Delivery of Services in Multi-Layer Multi-Domain Optical Network Infrastructures

Georgakilas, Konstantinos

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

Delivery of Services in Multi-Layer Multi-Domain Optical Network Infrastructures

Konstantinos N. Georgakilas

Department of Electronic Systems
Aalborg University

Denmark, 2012

Thesis title:	Delivery of Services in Multi-Layer Multi-Domain Optical Network Infrastructures
PhD Student:	Konstantinos N. Georgakilas
Academic Supervisor:	Jens Myrup Pedersen, Associate Professor, Department of Electronic Systems, Aalborg University
Co-supervisor:	Anna Tzanakaki, Associate Professor, Athens Information Technology
Co-supervisor:	Ole B. Madsen, Professor Emeritus, Aalborg University
Defense Date:	January 7, 2013
Assessment Committee:	Professor George N. Rouskas, Dept. of Computer Science, North Carolina State University, USA. Assistant Professor Christina Politi, Dept. of Telecommunications Science & Technology, University of Peloponnese, Greece. Associate Professor Henrik Schioler, Dept. of Electronic Systems, Aalborg University, Denmark.

List of published papers:

1 K. N. Georgakilas, A. Tzanakaki, M. Anastasopoulos, and J. M. Pedersen, "Converged Optical Network and Data Center Virtual Infrastructure Planning," J. Opt. Commun. Netw., vol. 4, no. 9, pp. 681–691, 2012.

2 K. N. Georgakilas, K. Katrinis, A. Tzanakaki, and O. B. Madsen, "Performance Evaluation of Impairment-Aware Routing Under Single- and Double-Link Failures," J. Opt. Commun. Netw., vol. 2, no. 8, p. 633, 2010.

3 Georgakilas, K.; Tzanakaki, A.; , "The impact of optical wavelength conversion on the energy efficiency of resilient WDM optical networks," Transparent Optical Networks (ICTON), 2011 13th International Conference on , vol., no., pp.1-4, 26-30 June 2011.

4 Georgakilas, K.N.; Katrinis, K.; Tzanakaki, A.; Madsen, O.B.; , "Impact of dual-link failures on impairment-aware routed networks," Transparent Optical Networks (ICTON), 2010 12th International Conference on , vol., no., pp.1-4, June 27 2010-July 1 2010.

Submitted Paper:

5 K. N. Georgakilas, M. P. Anastasopoulos, J. M. Pedersen and A. Tzanakaki, "Stochastic Optimization for Deployment of Correlated Cloud-Based Services over Optical Networks ,", IEEE Journal on Selected Areas in Communications, Special Issue on Networking Challenges in Cloud Computing Systems and Applications, submitted December 2012.

Other Publications:

1 Landi, G.; Ciulli, N.; Buysse, J.; **Georgakilas, K.**; Anastasopoulos, M.; Tzanakaki, A.; Develder, C.; Escalona, E.; Parniewicz, D.; Binczewski, A.; Belter, B.; , "A Network Control Plane architecture for on-demand co-provisioning of optical network and IT services," Future Network & Mobile Summit (FutureNetw), 2012 , vol., no., pp.1-8, 4-6 July 2012

2 Tzanakaki, A.; Anastasopoulos, M.P.; **Georgakilas, K.**; Garcia-Espin, J.A.; Riera, J.F.; Figuerola, S.; Ghijsen, M.; Demchenko, Y.; de Laat, C.T.A.M.; Vicat-Blanc, P.; Soudan, S.; Anhalt, F.; Peng, S.; Escalona, E.; Nejabati, R.; Simeonidou, D.; , "Virtual infrastructure planning: The GEYSERS approach," Future Network & Mobile Summit (FutureNetw), 2012 , vol., no., pp.1-9, 4-6 July 2012

3 Shuping Peng; Nejabati, R.; Escalona, E.; Simeonidou, D.; Tzanakaki, A.; Anastasopoulos, M.; **Georgakilas, K.**; , "Optimized PLI-aware virtual optical network composition," Transparent Optical Networks (ICTON), 2012 14th International Conference on , vol., no., pp.1-4, 2-5 July 2012

4 Anastasopoulos, M.P.; Tzanakaki, A.; **Georgakilas, K.**; , "Stochastic virtual infrastructure planning in elastic cloud deploying optical networking," Optical Fiber Communication Conference and Exposition (OFC/NFOEC), 2012 and the National Fiber Optic Engineers Conference , vol., no., pp.1-3, 4-8 March 2012

5 Buysse, J.; **Georgakilas, K.**; Tzanakaki, A.; De Leenheer, M.; Dhoedt, B.; Develder, C.; Demeester, P.; , "Calculating the Minimum Bounds of Energy Consumption for Cloud Networks," Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on , vol., no., pp.1-7, July 31 2011-Aug. 4 2011

6 Tzanakaki, A.; Anastasopoulos, M.P.; Georgakilas, K.; Simeonidou, D.; , "Energy aware planning of multiple virtual infrastructures over converged optical network and IT physical resources," Optical Communication (ECOC), 2011 37th European Conference and Exhibition on , vol., no., pp.1-3, 18-22 Sept. 2011

7 Tzanakaki, A.; Anastasopoulos, M.; Georgakilas, K.; Buysse, J.; De Leenheer, M.; Develder, C.; Peng, S.; Nejabati, R.; Escalona, E.; Simeonidou, D.; Ciulli, N.;

Landi, G.; Brogle, M.; Manfredi, A.; Lopez, E.; Riera, J.F.; Garcia-Espin, J.A.; Donadio, P.; Parladori, G.; Jimenez, J.; , "Energy Efficiency in integrated IT and optical network infrastructures: The GEYSERS approach," Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on , vol., no., pp.343-348, 10-15 April 2011

8 M. Anastasopoulos, A. Tzanakaki,**K. Georgakilas**, and D. Simeonidou, "Energy aware planning of multiple virtual infrastructures over converged optical network and IT physical resources," Opt. Express 19, B503-B508 (2011).

9 M. Anastasopoulos, A. Tzanakaki, and **K. Georgakilas**, "Evolutionary optimization for energy efficient service provisioning in IT and optical network infrastructures," Opt. Express 19, B496-B502 (2011).

10 Anastasopoulos, M.P.; Tzanakaki, A.; **Georgakilas, K.**; , "Virtual Infrastructure Planning in Elastic Cloud Deploying Optical Networking," Cloud Computing Technology and Science (CloudCom), 2011 IEEE Third International Conference on , vol., no., pp.685-689, Nov. 29 2011-Dec. 1 2011

11 Anastasopoulos, M.P.; **Georgakilas**, K.; Tzanakaki, A.; , "Evolutionary optimization for energy efficient service provisioning in IT and optical network infrastructures," Optical Communication (ECOC), 2011 37th European Conference and Exhibition on , vol., no., pp.1-3, 18-22 Sept. 2011

12 Tzanakaki, A.; Anastasopoulos, M.P.; **Georgakilas, K.**; Simeonidou, D.; , "Energy aware planning of multiple virtual infrastructures over converged optical network and IT physical resources," Optical Communication (ECOC), 2011 37th European Conference and Exhibition on , vol., no., pp.1-3, 18-22 Sept. 2011

13 Buysse, J.; **Georgakilas**, K.; Tzanakaki, A.; De Leenheer, M.; Dhoedt, B.; Develder, C.; Demeester, P.; , "Calculating the Minimum Bounds of Energy Consumption for Cloud Networks," Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on , vol., no., pp.1-7, July 31 2011-Aug. 4 2011

14 Gutierrez, J.M.; Riaz, T.; Pedersen, J.M.; Madsen, O.B.; **Georgakilas, K.**; Katrinis, K.; Tzanakaki, A.; , "Single and Dual physical link failures stability effect on

degree three WDM networks," Communications (APCC), 2010 16th Asia-Pacific Conference on , vol., no., pp.89-94, Oct. 31 2010-Nov. 3 2010

15 Gutierrez, J.M.; Katrinis, K.; **Georgakilas, K.**; Tzanakaki, A.; Madsen, O.B.; , "Increasing the cost-constrained availability of WDM networks with degree-3 structured topologies," Transparent Optical Networks (ICTON), 2010 12th International Conference on , vol., no., pp.1-4, June 27 2010-July 1 2010

16 Tzanakaki, A.; **Georgakilas, K.**; Katrinis, K.; Wosinska, L.; Jirattigalachote, A.; Monti, P.; , "Network performance improvement in survivable WDM networks considering physical layer constraints," Transparent Optical Networks, 2009. ICTON '09. 11th International Conference on , vol., no., pp.1-4, June 28 2009-July 2 2009

Abstract

The diffusion of Internet in society both for simple users and enterprises through the rise of applications and business models has resulted in a vast volume of traffic that needs to be transferred across the globe in a fast, reliable and secure way. Millions of fixed and mobile users produce and consume information, especially due to the recent increase of smart phones and tablets technology on the one hand and the fast growth of services like video, high definition TV and social networking on the other. Speed, reliability and minimal delay have been some of the key requirements for service delivery across the years of Information and Communications Technology (ICT) evolution. Current and future services take these requirements to a new level, not only because of the huge number of users and end-devices, but also due to the special requirements they have. To support these, new types of infrastructures have been developed such as interconnected Data Centers through high speed networks offering cloud-based services. Optical networks through recent technology advancements offer a very promising solution to support not only traditional telecommunications applications but all the new applications emerging in the context of cloud computing, due to the huge bandwidth they offer, the long transmission distances and several other characteristics including flexibility, cost and energy efficiency.

Apart from the evolution of network equipment technology offering among others improved physical layer performance across long distances, regeneration capabilities, low energy consumption and fast recovery, significant effort is required on the track of network design and operation. The work presented in this thesis concentrates in optical networking in support of traditional telecommunications and more advanced cloud computing services. More specifically it addresses some key optical network design and provisioning problems and aims at identifying efficient solutions to mitigate them. A core optical network able to efficiently accommodate traffic with special characteristics and requirements is the basis of every problem we address. In this context, physical layer impairments and their consideration in route selection, resilience against link failures, energy efficiency and a special case of service deployment in converged optical networks and Data Center (DC) infrastructures are the four main areas that this work focuses on through the formulation and solution of associated problems.

Core optical networks are commonly based on mesh topologies that accommodate

highly aggregated traffic and span ultra long-haul distances to interconnect major cities across the globe. Optical links though suffer from transmission and switching impairments that degrade the quality of the optical signal traversing the network. This can lead either to very low quality of service or to high connection blocking rates if a lower value of the path Bit Error Rate is considered acceptable. This effect is more prevalent when path protection mechanisms are present to mitigate the impact of link failures. Dedicated and shared protection paths are commonly longer than primary paths, since the shortest paths are typically used to establish the primary paths of connection requests and thus physical impairments need to be carefully considered. In this context, part of this thesis addresses both physical impairments and link failures (single and dual) in terms of a performance evaluation study under dynamic traffic arrivals. Physical layer impairment aware routing is compared to classical shortest path routing under the shared backup path protection scheme and several metrics are used to evaluate the performance of the proposed approach in the presence of dual link failures.

Path protection mechanisms in core optical networks offer 100% restorability against single link failures and very good performance under dual-link failures. However, the excess protection capacity requires more equipment to be used, even in the case where sharing of backup links is an option. Part of the work presented in this thesis quantifies through dimensioning problems the impact of dedicated and shared path protection schemes on the total power consumption of the network equipment and investigates the energy savings offered by the use of all-optical against Optical-Electrical-Optical (OEO) wavelength conversion.

Cloud computing supported by interconnected Data Centers and the adoption of infrastructure virtualization has recently widely penetrated in Information and Communication Technology (ICT). Optical networking is a very promising candidate solution for the interconnection of Data Centers, both internally (intra-Data Center) and externally (inter-Data Center). Focusing on inter-Data Center communication as part of a converged optical network and Data Center Infrastructure, this thesis also examines how the required virtual infrastructures (VIs) can be designed to facilitate sharing of the physical infrastructure among different service providers. In particular, we formulate and solve network design problems that take into account both network and DC resources. We compare two objectives, one minimizing the joint power consumption of network and DC resources and one minimizing the network resources used. The goal of

this comparison is to identify suitable design objectives, trade-offs and trends for realistic VI request scenarios and a variety of traffic loading conditions. Moreover, we study the impact of the design objectives on the resulting virtual topologies and their performance under dynamic traffic.

Taking a more service-driven view of planning a converged optical network and DC infrastructure, we investigate how the correlation patterns among cloud-based services that is not usually taken into account impacts their deployment across DCs and the network itself. A stochastic optimization approach is used to illustrate how the service correlation consideration as a random process provides a better placement both in terms of DC and network resources. Appropriate deterministic problems are also solved to illustrate the impact of service correlation in the total cost when this is not taken into account during the planning process.

Dansk resumé

Udbredelsen af Internettet i samfundet blandt både almendelinge og mere avancerede brugere og virksomheder har pga. antallet af applikationer og forretningsmodeller forårsaget en kolossal forøgelse af trafikken, som skal føres verden rundt på en hurtig, pålidelig og sikker måde. Millioner af faste og mobile brugere producerer og forbruger information, især på grund af stigningen indenfor smartphones og tablets, og den hurtige vækst indenfor tjenesteydelser som f.eks. video, højopløsnings-TV, og sociale netværk. Hastighed, pålidelighed og tidslighed har været nogle af de vigtigste krav indenfor udviklingen af levering af informationsteknologiske tjenesteydelser igennem årene. Nuværende og fremtidige tjenesteydelser øger disse krav til et nyt niveau, ikke kun pga. det store antal brugere og apparater, men også pga. de specielle krav der stilles. For at understøtte disse krav har man udviklet nye typer af infrastrukturer såsom sammenkoblede datacentre (DC) via højhastighedsnetværk, som tilbyder cloud-baserede tjenesteydelser. Optiske netværk, som er udviklet på baggrund af de nyeste teknologiske fremskidt, fremstår som en meget lovende løsning som udover traditionelle applikationer indenfor telekommunikation understøtter alle nye applikationer som er dukket op ifm. cloud computing. Dette skyldes bl.a. store båndbredder, lange transmissionsafstande, fleksibilitet, lav pris, energieffektivitet og mange andre egenskaber.

Den teknologiske udviklingen indenfor netværk og udstyr betyder bl.a. forbedret ydeevne på det fysiskelag – over lange afstande – herunder mulighed for genetablering af forbindelser ved fejl, lavt energiforbrug og hurtig fejlretning. Dog er det afgørende med hensigtsmæssigt design og drift. Denne afhandling har fokus på optiske netværks understøttelse af traditional telekommunikation såvel som mere avancerede cloud computing tjenester. Mere specifikt addresseres de væsentligste problemer indenfor netværksdesign og udrulgning med henblik på at identificere effektive løsninger der mindsker problemer. Et optiske core netværk, som er i stand til effektivt at understøtte al trafik med specielle karakteristikker og krav er grundlaget for alle de problemer vi forsøger at løse.

I den forbindelse er problemer på det fysiske lag og deres indvirkning på rutevalg, robusthed ved linkfejl, energieffektivitet og et særtilfælde af serviceetablering i konvergerede optiske netværk og DC infrastrukturer de fire hoveområde dette arbejde fokuserer på gennem formulering af problemer og løsninger heraf.

Core optiske netværk er som regel baseret på mesh-topologier, som understøtter store mænger akkumuleret trafik og forbinder større byer rundt i verden over meget lange distance. Optiske forbindelser lider dog under at det optiske signal forringes under transmission og switching, hvilket kan forringer kvaliteten af forbindelser, eller endog føre til blokering af forbindelser hvor der stilles krav om lave bit error rates. Dette ses især når man har mekanisker til path protection som kan mindske effekterne af fejl i forbindelsen. Dedikeret eller delt path protection medfører sædvanligvis længere veje end de primære veje, da man ved etablering af en forbindelse normalt vil vælge den korteste vej. Det er derfor vigtigt at de fysiske forringelser nøje overvejes. I den sammenhæng drejer en del af denne afhandling sig om både fysiske forringelser og linkfejl (enkelte og dobbelte), hvor der præsenteres evalueringer af performance under dynamiske ankomster af trafik. "Physical layer impairment aware routing" sammenlignes med klassiske shortest-path routing under antagelse a delte backup-veje, og der anvendes adskillige metrikker for at beskrive hvordan den foreslåede tilgang virker i tilfælde af dobbelte linkfejl.

Path protection mekanisker i core optiske netværk tilbyder 100% reetablering ved enkelte link-fejl, og rigtig god ydelse også ved dobbelte link-fejl. Men den ekstra beskyttelseskapacitet kræver anvendelse af mere udstyr, selv i tilfælde af at der er mulighed for fælles backup links, selv når backup links kan deles. Denne afhandling kvantificerer effekten af dedikere og fælles path protection metoder i forhold til netværksudstyrets totale strømforbrug, og undersøger hvor meget energi der kan spares ved hjælp af fuldt optiske (i modsætning til optiske-elektrisk-optisk) konvertering af bølgelængder.

Cloud Computing som understøttes af sammenkoblede DC, og indførelsen af virtualisering af infrastrukturen, har de senere år opnået stor udbredelse indenfor ICT. Optiske netværk er en oplagt kandidat til at forbinde DC, både internt (intra-data center) og eksternt (inter-data center). Denne afhandling fokuserer på de eksterne forbindelser som en del af konvergerede optiske netværk og DC infrastruktur, og undersøger bl.a. hvordan de nødvendige virtuelle infrastrukturer (VI) kan designes til at facilitere deling af den fysiske infrastruktur mellem forskellige serviceudbydere.

Vi vil især formulere og løse netværk design problemer, som tager hensyn til både ressourcerne i både netværk og DC. Vi sammenligner to mål: Den ene minimerer det fælles strømforbrug i netværket og DC, den anden minimerer de netværksressourcer der bliver brugt. Målet med denne sammenligning er, at identificere egnede designmuligheder, trade-offs og tendenser for realistiske scenarier for VI efterspørgsler og en

række betingelser i fht trafikbelastninger. Endvidere studerer vi virkningen af de forskellige design på de deraf følgende virtuelle topologier og deres ydeevne under dynamisk trafik.

Ud fra et mere service-drevet synspunkt på planlægningen af et konvergeret optisk netværk og DC infrastruktur, undersøger vi, hvordan sammenhængende mønstre i cloud-baserede tjenesteydelser, som der normalt ikke tages hensyn til, påvirker deres indsættelse i DC og i selve netværket. Der anvendes en stokastisk optimering for at illustrere hvordan service sammenhængen, som en tilfældig proces, giver en bedre placering både ift. DC og netværksressourcer. Egnede deterministiske problemer løses også, for at illustrere virkningen af service sammenhængen i de totale omkostninger, når der ikke tages hensyn hertil i planlægningsprocessen.

Acknowledgments

I would like to thank my supervisor Anna Tzanakaki for her invaluable mentorship, inspiration and support throughout the process of this work, for introducing me to the field and guiding me towards the achievement of my goals based on scientific and human values; my supervisors Jens M. Pedersen and Ole B. Madsen for their continuous support and advice; also, Kostas Katrinis for his guidance and collaboration during the initial years of this work; and my latest collaborator Markos Anastasopoulos for his stimulating ideas and support.

Finally and most of all, I would like to thank my parents and big brother, without whom I would not be able to complete this work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since the mid-90 's and until today, the Internet has evolved to a complex network of networks that interconnects billions of users. Simple information exchange between users has transformed to the use of applications such as video, social networking, high definition TV, gaming and more through millions or billions of devices [1] such as laptops, smart phones, tablets and sensors. Users both produce and consume information that is distributed between multiple points across the globe. The estimated traffic growth that is of the order of 50% per year [2] and the explosion of connected devices creates a tidal wave of structured and unstructured data, also known as "big data"[3]. Apart from new applications and business models generated, big data has led to new concepts of handling processing, storage and provisioning of services in a fast, reliable and secure way. The Cloud [4] paradigm is the most recent and most promising approach of such technology that takes the challenge of data management in terms of volume, variety, velocity and veracity [3] to a new level through advanced Data Center and virtualization technology. This evolution generates the need for a network that can accommodate huge amounts of traffic with very low latency across long distances and with demanding quality of service requirements such as reliability, security, synchronization and minimum delay. Optical networks played an important role in communication networks during the last thirty years and have evolved as the most promising technology to support this need.

The evolution of optical network technology is translated through radical advancements in transmission capacity, signal bandwidth and transmission distance and has been supported by three technological advancements: a) Time Division Multiplexing (TDM), b) optical amplification and Dense Wavelength Division Multiplexing (DWDM) and c) alternative modulation formats and coherent detection [2]. Coherent detection in particular along with digital signal processing is a promising transport technology since it aims to address several important issues and offer characteristics such as higher signal to noise ratio (SNR) and compensation for linear distortion and polarization-mode dispersion in the optical path (Figure 1.1) [2]. This evolution of optical transport technology illustrates some of the most important problems and challenges that need to be addressed such as physical layer impairments, fault tolerance and energy efficiency.

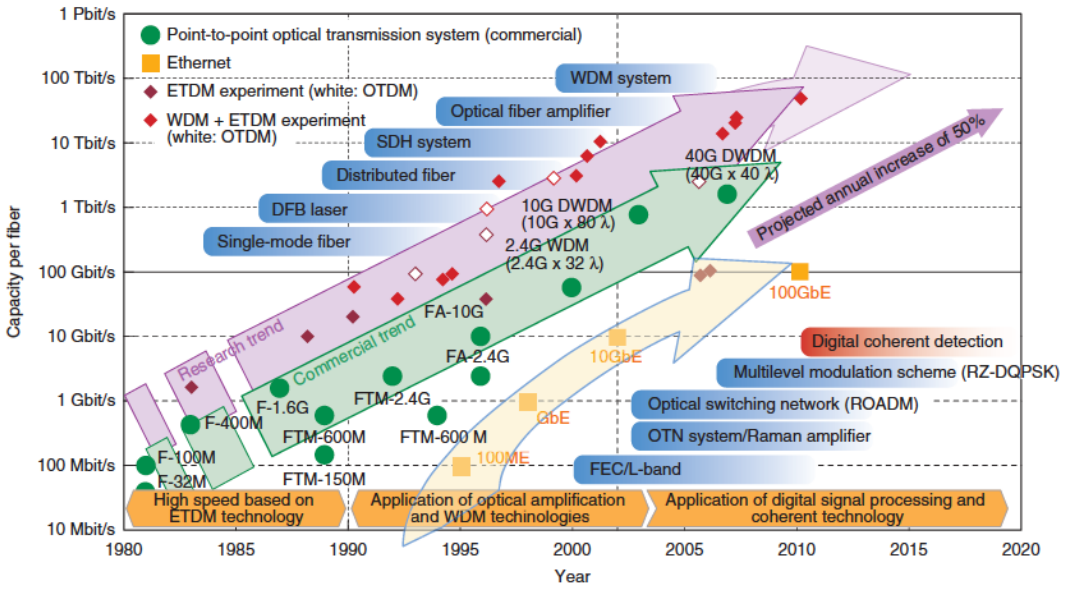

Figure 1.1 – Optical transport technology trends [2]

Core optical networks are usually mesh topologies and interconnected rings with nodes in major cities that interconnect metro and access networks. The traffic coming through each core node is usually destined to all other core nodes due to high aggregation, resulting in highly connected nodes. Moreover, core networks span very long distances to interconnect cities in a country, continent or the globe and in the case of transparent

optical networks without intermediate technology like repeaters or regenerators. The effect of this characteristic is two-fold: a) physical layer impairments on the fiber links are much more prevalent due to ultra long-haul distances and b) possible failures of node or link equipment will result in the disruption of a vast amount of traffic and will probably take days or weeks to be repaired.

Although technology of transmission and switching equipment has lowered the power consumption of the optical network, the current and especially the estimated traffic growth requires multiple fibers, large scale optical cross connects and respective control equipment to be supported. Power consumption for telecommunication networks is estimated to increase by 27% in 2020 [5], and thus converged optical networks and Data Center (DC) infrastructures are expected to have an important impact on power consumption, making energy efficiency a goal not only for manufacturing but also for the design and operation of converged network and Cloud infrastructures.

The proliferation of Cloud computing and the penetration of cloud-based services changes the characteristics and requirements of the design and operation for the supporting network and Information Technology (IT) infrastructures. Apart from joint consideration of network and IT resources in all phases of the infrastructure planning and service provisioning, new paradigms of cloud-based services delivery such as Virtual Machines (VMs) require additional details to be taken into account. An important characteristic of the services offered by distributed Data Centers is the inter-dependence between them and the additional background communication that this requires. As shown in the literature recently [6], cloud-based services demonstrate some correlation patterns that require a pair of correlated services offered in distinct DCs to exchange data before the client requesting one of them is served. If this is taken into account in the way that VMs are deployed across the multiple DCs, then a significant benefit on the network utilization and the service delivery as a whole can be achieved.

In this work we formulate and solve service provisioning and network design problems aiming to illustrate the aforementioned problems and challenges and demonstrate efficient solutions towards core optical networks that can efficiently and effectively support

the increasing volume of traffic. We thus tackle four main areas of interest related to core optical networking: a) physical layer impairments and resilence, b) resilience and energy efficiency, c) energy efficiency of converged optical network and DC infrastructures and d) deployment of correlated cloud-based services over optical networks.

1.1 Optical Network Infrastructure Technologies

Wavelength Division Multiplexing (WDM) is the key transmission technology that has enabled the transfer of huge traffic volumes across the globe. " *With WDM, multiple channels, each with a different signal, are transmitted at distinct wavelengths over a single optical fiber. In this way, the capacity of a single fiber is upgraded to a multiple of its original capacity. The capacity of an optical fiber is, thus, no longer limited to the bit rate of a single TDM signal, but by the number of wavelengths supported by the WDM system.* "[7]. Each signal is thus transmitted on its own dedicated bandwidth defined by the wavelength and all signals are transmitted at the same time. Dense WDM (DWDM) achieves closer spacing of wavelengths (frequencies) offering a greater number of wavelengths per fiber and thus greater bandwidth [8].

With the recent technology evolution in the optical communications domain, the WDM transport layer evolved from simple point-to-point transmission links into elaborate network architectures providing similar functionality to the electronic layer, with improved features, higher manageability and lower complexity and cost [9], [10], [11]. Integrated WDM networks performing switching and routing are deployed in order to economically support the required functionalities. In such network scenarios, high capacity optical paths (lightpaths) are set in the transport layer forming connections between discrete points of the network topology. These can be identified to be reconfigurable Optical Add/Drop Multiplexers (OADMs) and Optical Cross-Connect (OXC) nodes performing traffic engineering and management of the optical bandwidth [12], [13]. More specifically they support handling of the incoming signals at the appropriate granularity

level to enable efficient routing of the traffic demands satisfying the service level requirements including network survivability and security and accommodate network expansion, traffic growth and churn. This type of nodes can offer functionalities such as dynamic service provisioning and bandwidth on demand while advanced designs equipped with the required hardware and software are able to also support enhanced network features and new services. These functions are facilitated through the application of the Automatically Switched Otical Network (ASON) and the Generalized Multi-Protocol Label Switching (GMPLS) as the standardized and common control plane suite used in this type of networks.

Lightpaths and wavelength routed networks [14] are the basic elements of wide area network architectures that constitute the core of the global network hierarchy (Figure 1.2]) as presented in [8]. Typically, long-haul architectures consist of OXCs interconnected through fiber links in mesh topologies (or rings) [14] and provide connectivity for access and metro networks through switching and routing functions.

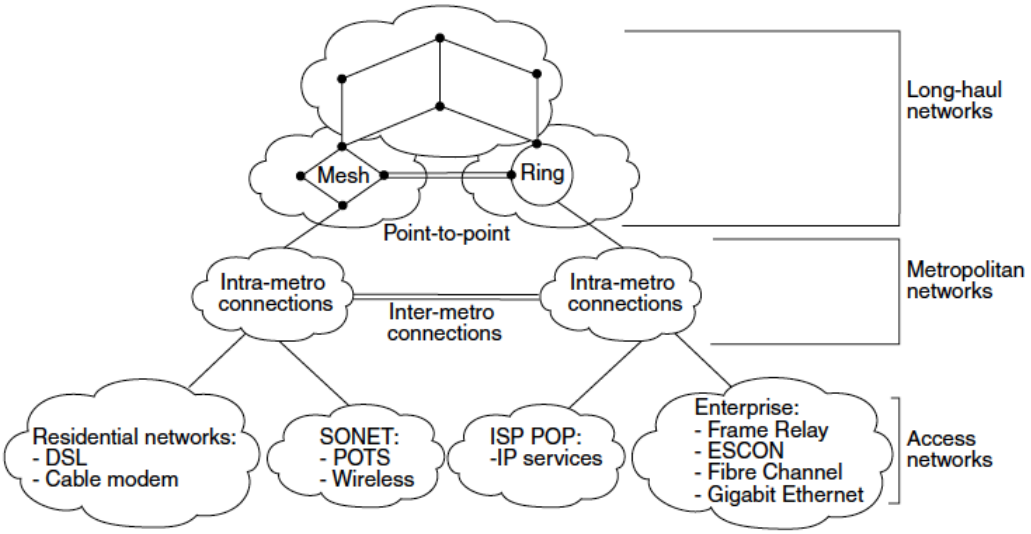

Figure 1.2 – Network Hierarchy [8]

Optical network technologies are classified into optical transmission, switching and control technologies depending on their role and functionality in the network. A relevant

taxonomy diagram is illustrated in Figure 1.3

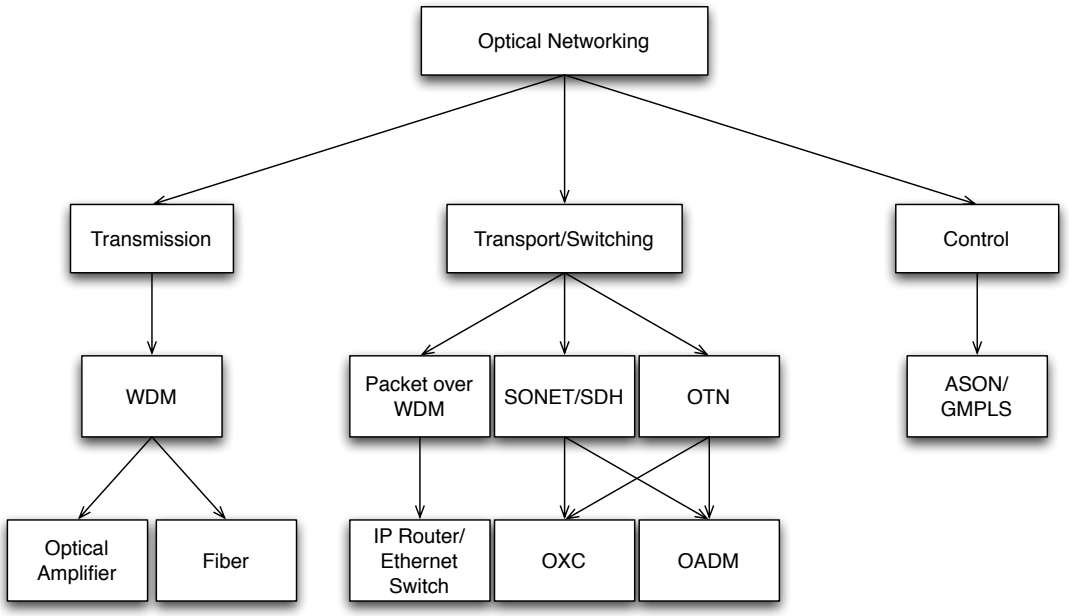

Figure 1.3 – Optical Network Taxonomy

The main architectures that have been proposed to date for core transport and switching are based on switched Synchronous Optical Network (SONET)/Synchronous Digital Hierarchy (SDH), switched Optical Transport Network (OTN) and IP over DWDM technologies. Their main characteristics are presented below.

SONET is the current transmission and multiplexing standard for high-speed signals within the carrier infrastructure in North America while SDH has been adopted in Europe and Japan and for most submarine links. For SONET the basic signal rate is 51.84 Mb/s, called the synchronous transport level-1 (STS-1). Higher rate signals (STS-N) can be obtained by interleaving the bytes from N frame aligned STS-1s. For SDH the basic rate is 155 Mb/s and is called STM-1 (synchronous transport module-1), which is higher than the basic SONET bit rate.

Switched OTN is an optical transport standard defined by the ITU-T G.709 standards committee that contains definitions for payload encapsulation, OAM overhead,

FEC and multiplexing hierarchy. It includes some of the SONET/SDH benefits (resilience and manageability), fault detection, communication channels and multiplexing hierarchy. It is designed to be a multi-user transport container for any type of service (TDM, packet) and provides end-to-end optical transport transparency of customer traffic. Switched OTN is widely deployed for transport within long-haul networks, mainly because the longer distances of optical transmission enabled by the inherent forward error correction (FEC) mechanism [15].

IP over DWDM has been proposed as an alternative to SONET/SDH for IP packet transmission over optical fiber networks. The all-optical transport layer is more cost efficient (simplification of the network layers) and maintains high data rates. Benefits of this solution are related to faster path provisioning. However, several disadvantages arise from the fact that router ports are expensive compared to switch or transmission cost. In addition, inherent scalability issues associated with the IP router technology as well as the very high energy consumption levels associated with this type of equipment when compared to their optical technology counterparts, may introduce serious drawbacks regarding their suitability for a sustainable Future Internet solution.

Gigabit and 10 Gigabit Ethernet is based on a bus architecture where all the nodes are connected to a single bus. Gigabit Ethernet is an extension of the same standard to 1 Gb/s. It operates over both copper and fiber interfaces. Gigabit over fiber is becoming a popular choice in metro networks to interconnect multiple enterprise networks.

The transport rates and the corresponding technology are another choice that the operators have to take along with the deployment of the appropriate transport and switching technologies. The available options are 1) the continuity of 10G capacity placement and 2) the upgrade to 40/100G. The former solution benefits from the price reduction of 10G technology that is expected to continue and the existence of the corresponding network standards for more than a decade. However, with traffic growth estimated to be in the order of 50% per year and the operational complexity and inefficiency that the placement of more 10G capacity acquires, the second option seems as a more appropriate solution. 40G has already moved from early adoption to massive deployment from

several operators around the globe including AT&T, Verizon, DT, China Telecom and others. Moreover, 40GE and 100GE transport over 40G and 100G networks is already fully standardized and some initial deployments of commercial 100G already exist [15].

Transparency in optical networking refers to the ability to modulate and transmit any kind of payload on the optical channel, independent of its bit-rate and format (framing, line-coding, power level, etc). Transparency implies that a specific lightpath is assigned between each origin and destination node pair without any optical-electronic-optical (OEO) conversion at any intermediate node. In the general case, transparent optical networks provide reduced operational costs associated with their inherent energy efficiency and small footprint but suffer from the physical layer impairments associated with the optical transmission and switching of the data channels. In addition, they do not inherently support wavelength conversion capability and signal monitoring functions. However, wavelength conversion capabilities can be introduced through the use of transparent optical wavelength converters based on all optical technologies [16].

Opaque networks are on the other hand based on nodes equipped with OEO technologies. These can more specifically be either receiver/transmitter pairs associated with an optical switching fabric in which case the finest granularity that the network supports is that of the wavelength, or receiver/transmitter pairs associated with an electrical switching fabric in which case the switching granularity supported by the network could also be sub-wavelength. Typically, these networks inherently support wavelength conversion functionality and signal monitoring capabilities. However, they require higher energy consumption levels for their operation and occupy larger footprint compared to their transparent counterpart.

A practical solution that is commonly deployed with the aim to overcome the limitations of both transparent and opaque optical networks are translucent optical networks. These provide some limited level of transparency, based on what is commonly known as transparency islands, i.e., network parts that are fully transparent interconnected together through opaque network nodes including OEO signal conversion and therefore the associated technologies. This way, the overall network cost and power consumption

can be reduced but special network design considerations involving optimal equipment placement are required [17].

The presence or absence of wavelength conversion in the network plays a significant role in the way service provisioning is handled and the level of efficiency of the resource utilization achieved when provisioning the services. More specifically, in case of the absence of wavelength conversion, optical path assignment is performed by assigning the same wavelength across all links of a path through Wavelength Assignment (WA) algorithms. This is known as the wavelength-continuity constraint and is referred to as the pure Wavelength Path (WP) case. On the other hand, in case that wavelength conversion is available, the wavelength across the different links of a lightpath does not need to be the same, but it is assigned based on the associated bandwidth availability on a per link basis and it is referred to as the Virtual Wavelength Path (VWP) case. Several advantages arise from the presence of wavelength converters at the network nodes, such as lower blocking probability, lower complexity of algorithms, ease of provisioning and management of connections, lower capacity requirements and easier and feasible network design operations.

Through the GMPLS protocol suite, the Network Control Plane (NCP) enables the evolution from centralized to distributed control of access, metro, regional and long-haul networks. It operates over multiple vendor and operator environments and technologies including IP, Ethernet and optical networks and in a simplified view it has the role to dynamically setup connections across an optical transport network. The main benefits of an NCP are:

- Distributed and reactive traffic engineering, allowing network resources to be dynamically allocated to connections.
- Usage of specific control plane protocols rather than generalised network management protocols.
- Distributed and reactive restoration upon a network failure, taking into account current state of the transport network.

- Reusability of control plane protocols to handle different transport technologies under a common control framework.

1.2 Thesis Contributions and Outline

1.2.1 Impairment Aware Routing and Resilience

Chapter 2 of the work presented in this thesis focuses on the performance evaluation of an impairment-aware routing (IAR) scheme [18] in the presence of single and dual link failures. Impairment aware routing is adopted to identify both primary and backup paths and is compared with the case of applying impairment aware routing in the primary and conventional min-hop routing in the backup path computation. Protection against link failures is provided through a shared backup path protection scheme [19] enhanced with a reinforced sharing mechanism to take advantage of the increased resource utilization it offers. The novelty of this work lies in the presentation of a detailed study that systematically illustrates the effect of single and dual link failures in optical networks designed to offer single link failure resilience while taking into consideration the physical layer characteristics in the routing of both primary and protection paths.

1.2.2 Resilient Optical Networks and Energy Efficiency

Motivated by the increasing interest in energy-efficient networks, in Chapter 3 we compare the power consumption of two different WDM optical network architectures that both employ wavelength conversion at every intermediate node but use different technologies for carrying out the conversion. Conventional optical-electrical-optical (OEO) [20] and all-optical based on Semiconductor Optical Amplifiers (SOA) [21] are the two wavelength conversion technologies used. On top of these two architectures, we evaluate the impact of two path protection mechanisms in terms of their excess power consumption requirements compared to the unprotected case and also evaluate the savings that all-optical wavelength conversion technology offers. This is done by dimensioning the network through minimum-cost optimization [22] problems and post-calculating the

power consumption. The results have been produced assuming 10Gb/s and 40Gb/s per wavelength channel data rates and indicate that the use of all-optical wavelength conversion significantly assists in decreasing the overall network power consumption. In addition, we identify the SBPP (Shared Backup Path Protection) scheme as the most energy–efficient survivability solution.

1.2.3 Optical Network and Data Center Infrastructure Planning

Chapter 4 focuses on the design of Virtual Infrastructures (VIs) over a Physical Infrastructure (PI) taking into consideration jointly the network and DC resources. The VIs are slices of the PI comprising subsets of the optical WDM network and DC resources enabling sharing of the available physical resources among several virtual network operators and services [23]. The novelty of this work lies in the more realistic assumption of no global knowledge for the requests for all the VIs. We thus perform the planning of each VI in sequence according to the arrival order of the VI requests over the underlying PI that is already supporting previously established VIs. Through the design process, both the topology and required virtual resources are identified and mapped to the physical resources and the associated operating parameters. In this context, we compare two objectives, one minimizing the joint power consumption of network and DC resources and one minimizing the network resources used. The goal of this comparison is to identify suitable design objectives, tradeoffs and trends for realistic VI request scenarios and a variety of traffic loading conditions. Moreover, we study the impact of the design objectives on the resulting virtual topologies and their performance under dynamic traffic.

The planning problems are formulated with the objectives MinJointPower (MJP) and MinNetRes (MNR). Both objectives are evaluated over two network architectures: Virtual Wavelength Path (VWP), where full wavelength conversion is available across all network nodes and Wavelength Path (WP) [24], where wavelength continuity is a strict constraint. Finally, the respective VIs generated solving these planning problems are evaluated through online traffic provisioning simulations. The results demonstrate that

although the MJP objective achieves lower power consumption compared to the MNP as expected, the benefit decreases as the number of established VIs and the volume of demands supported increases. The performance comparison of the different planned VIs shows that the gain in the power consumption offered by the MNR objective introduces a penalty in the blocking performance. The presence of wavelength conversion in the network increases the overall power consumption but improves the blocking performance.

1.2.4 Optimal VM Deployment for Correlated Cloud-based Services

Motivated by the proliferation of cloud-based services and the need for optical networks to support the interconnection of geographically dispersed Data Centers (DCs), we present in the Chapter 5 of this thesis a novel piece of work in the context of optimal service deployment and network resource utilization. The unique characteristic of this work is the modeling of service cross-correlation as a stochastic process, based on the latest literature that captures the traffic characteristics for client-to-DC and DC-to-DC communication. We use the Sample Average Approximation method [25] to formulate the stochastic problems and compare them with appropriate deterministic problems that provide lower and upper bounds on the total DC and network infrastructure cost. The results identify the impact of correlation when it is deterministically considered on the one hand and the benefits of more accurately capturing this impact in a stochastic optimization [25] context.

Chapter 2

Impairment-Aware Routing Under Single and Dual Link Failures

Fault tolerance i.e., the resilience of the network to a wide range of failures is an aspect of major importance in WDM networks. Particularly link failures have been extensively studied, as optical fibers carry a large number of wavelength channels, each modulated at very high data rates exceeding 10, 40 or even 160 Gb/s. Optical networks are usually designed to offer survivability against single link failures with minimum additional capacity, while survivability in case of dual link failures has also been studied. Shared risk link groups (SRLGs) [26] and network maintenance [27] are two common reasons that may lead to a dual link failure. In this context, the impact of dual failures in the overall network performance has to be studied, as it can be of particular importance for networks requiring high degree of availability (e.g. carrier grade networks requiring higher than 99.999% availability). Therefore, several researchers have targeted this problem [28], [29], [30], [31], [32], [33] examining the effects that take place and proposing different solutions for WDM and IP-over-WDM networks.

Numerous classification schemes of link failures recovery mechanisms have been identified throughout the literature [26]. 1+1, 1:1 and M:N are well-known protection

schemes applied to point-to-point systems, whereas ring-based protection mechanisms apply to both ring and mesh topologies. Protection mechanisms can be classified according to the part of the network that the protection mechanism operates on, e.g. a link/span, a path or a segment. Although resilience is the general term under which all these schemes fall under, protection is a very common term used across schemes that should not though create confusion on the characteristics of each solution. For this reason, a usual classification deals with whether the protection path is computed a priori or in real time and usually (but not always) the respective terms used are protection and restoration. Specifically for mesh networks, the restoration schemes can be further classified as centralized and distributed. Furthermore, the sharing or not of protection capacity defines the recovery schemes as shared or dedicated, respectively [34]. Finally, single and multi-layer mechanisms have also been identified.

Regarding lightpath identification in optical networks, conventional routing and wavelength assignment (RWA) algorithms base the lightpath discovery only on network level considerations such as bandwidth availability, shortest distance etc, without any detailed physical layer consideration. However, optical networks suffer from transmission and switching impairments that degrade the quality of the optical signal traversing the network and for this reason several works address the consideration of physical layer impairments in the path computation process [35], [36], [37], [38], also referred to as Impairment Aware Routing (IAR) throughout this work. Typical physical layer impairments include power losses, chromatic dispersion (CD), polarization mode dispersion (PMD), polarization dependent loss (PDL) , amplifier spontaneous emission noise (ASE) and crosstalk (CT) regarding linear effects, and self-phase modulation (SFM), cross-phase modulation (XPM), four wave mixing (FWM) and stimulated Raman scattering (SRS) regarding non-linear effects [36]. Previous work has shown that the effect of optical network impairments on the signal quality is more profound in survivable WDM networks due to protection paths that are typically much longer than working paths [19], thus being highly susceptible to physical layer impairments.

In this chapter, we present a performance evaluation study of impairment-aware

routing under single and dual link failures for mesh optical WDM networks. Simulation results indicate that applying the impairment-aware routing scheme in the backup path computation provides significantly reduced blocking probability in the network compared to conventional minimum hop routing. The average connection loss rate is calculated in case of dual failures, illustrating the effect of dual failures on established and active connections in a network designed to handle single failures. In all scenarios under study the average connection loss rate remains below 2% for COST239 [39] and 3.5% for NSFNET [40] network topologies of the overall number of established connections. In addition to the protection scheme applied during the path provisioning phase, a restoration scheme that is activated in case of dual failures is also evaluated and its recovery capabilities are identified. Our simulation results show that a minimum of 60% restorability of the connections affected due to dual failures can be achieved for the COST239 network and 70% for the NSFNET. The results also demonstrate that the vulnerability of the network due to double link failures is not affected by the use of the IAR scheme for both topologies examined.

2.1 Related Work

2.1.1 Dual Link Failures

The general topic of resilience in optical networks has been extensively addressed in the literature to date. References [26] and [34] provide a detailed description of resilience in optical networks and the relevant issues. In [41] Clouqueur et al. focus on the availability of paths in span-restorable mesh networks that are designed to provide full single link failure restorability, under dual link failures, using span restoration. They provide a thorough theoretical analysis regarding end-to-end path availability, link unavailability and network average dual failure restorability. Moreover, they provide experimental results for dual failure restorability over several network topologies designed for full single failure restorability. In [42] the authors develop a theoretical analysis of the probability of multiple failures in the network, illustrating that all well-known resilience schemes are

vulnerable to multiple failures. In [43] Lumetta et al. present a classification of dual-link failures in WDM mesh networks depending on how they occur and identify the types of recovery mechanisms that each failure class affects. Their results illustrate that reconfiguration significantly increases the dual failure restorability, while shared path protection with reconfiguration achieves similar results compared to dedicated path protection and dedicated and shared link protection algorithms. The authors in [28] formulate an ILP problem that optimizes the network capacity for a path-based, dual-link failure recovery scheme, also identifying the scenarios where the sharing of wavelengths between backup paths does not violate full restorability. By comparing shared protection with dedicated protection, they demonstrate that the former utilizes 22%-38% less capacity than the latter, whereas the average capacity used for protection paths is 14% higher in the shared scheme. In [29] the authors present a new rule for sharing backup path resources in a shared path protection scheme in order to save redundant resources and also a new routing approach in order to find feasible solutions with three link-disjoint paths. Assuming full wavelength conversion capability across the network, they compare shared link protection, shared-path protection and the proposed enhanced shared backup path protection. Their results indicate significant resource savings and blocking probability performance improvement. In [30], SBPP is extensively studied and the authors produce a minimum capacity design for a given demand matrix, resulting in 70-80% dual failure restorability. Moreover, these results are extended by differentiating the sharing limit and conclude that a limit of 3 or less has to be used in order to achieve some real benefit in the trade-off between utilization and dual failure restorability. In [44], the impact of resource sharability on double link failure restorability is studied. They demonstrate that limiting the sharing degree (or sharing index) also limits the number of connections that need to be re-provisioned after a first link failure in order to be protected by a possible second failure. However, lower sharing also impacts the availability of resources for re-provisioning, thus an important trade-off is identified. In [45], Schupke et al. study the "working and protection capacity efficiency and the dual failure restorability for protected WDM networks with full wavelength conversion". The protection schemes studied

are dedicated path protection, shared path protection and path rerouting. The problems addressed are formulated as mixed integer linear programs that maximize dual failure restorability and optimize capacity in a second step, whereas they also present heuristics for path rerouting and larger problems. Their results illustrate that connectivity is more important than the total capacity for restorability and that rerouting outperforms both dedicated and shared path protection by 10-15% in terms of restorability.

2.1.2 Optical Physical Layer Impairments and Routing

In [46], Mukherjee et al. introduce the notions of opaque, translucent and transparent networks. Focusing on all-optical transparent networks, the authors differentiate the blocking of a lightpath establishment to what is referred to as physical layer blocking and to the well-known network layer blocking. The former occurs due to very high bit error rate (BER) of the lightpath caused by the accumulation of physical layer impairments across the multiple optical links that the lightpath traverses. The latter occurs due to unavailability of resources i.e., fiber bandwidth or wavelength. After referring to the most common linear and non-linear physical layer impairments and how these are connected to the increasing data rates, the authors provide a basic classification of routing approaches based on physical layer impairments. Finally, they demonstrate a blocking benefit of 30% for an impairment-aware RWA algorithm compared to a traditional approach. In [47] the authors provide the motivation for working on impairment aware routing algorithms and refer to the challenge of modeling the heterogeneous aspects of networking and physical layer considerations. Extending previous work on the subject, they present a new approach that takes into account both linear and non linear effects through the use of the Q-factor penalty assigned to each link. Their results demonstrate significant blocking performance improvement for an impairment aware shortest path approach compared to a conventional shortest path algorithm. [48] presents an "online constraint-based routing algorithm"that also takes into account both linear and non-linear effects of the optical layer. Both offline and online versions of the algorithm

are presented. The results demonstrate reduced blocking probability compared to typical shortest path routing with a trade-off in computational complexity. [19] presents the consideration of physical layer impairments in the process of performing RWA for a survivable optical WDM network. In more detail, the authors present an impairment aware routing and wavelength assignment approach applied on top of a backup multiplexing recovery mechanisms that is designed to provide full protection against single link failures. First, they examine the performance of three different wavelength assignment algorithms for the backup path, namely First Fit, Last Fit and Random Fit, while First Fit is always used for the primary path. The results demonstrate that Last Fit outperforms Random Fit and provides better blocking results even compared to First Fit for higher loading conditions. Moreover, comparing IAR and shortest path for the primary path in terms of blocking probability as a function of dispersion mapping, wide regions of optimal performance are identified when the impairment aware approach is used.

2.2 Simulation Model

The simulation model used in the framework of this chapter solves the online RWA/resilience problem, where traffic requests arrive and get served sequentially, without knowledge of future incoming requests. Since wavelength conversion is not considered available in any network node, the wavelength continuity constraint applies across the network. Incoming requests adhere to a Poisson arrival process with exponentially distributed time duration. Origin and destination nodes are randomly selected among the network nodes with a uniform distribution. 100% single link failure recovery is provided through a Shared Backup Path Protection (SBPP) recovery scheme that is already shown to demonstrate excellent utilization performance. In that sense, each connection request is established only after a primary and a link-disjoint backup path are identified from the RWA scheme applied.

The physical bandwidth of each link l is divided into the following three parts: A_l, B_l

and R_l [19], as illustrated in Figure 2.1. A_l represents the total bandwidth dedicated to primary paths carried by link l and is not allowed to be shared. B_l is the total bandwidth occupied by backup paths on link l and sharing by backup paths whose corresponding primary paths are link-disjoint is allowed. The residual bandwidth R_l is the difference between the total bandwidth of link l and the bandwidth consumed by primary and backup paths on that link $(A_l + B_l)$. No restriction regarding the percentage of primary and protection capacity on each link applies.

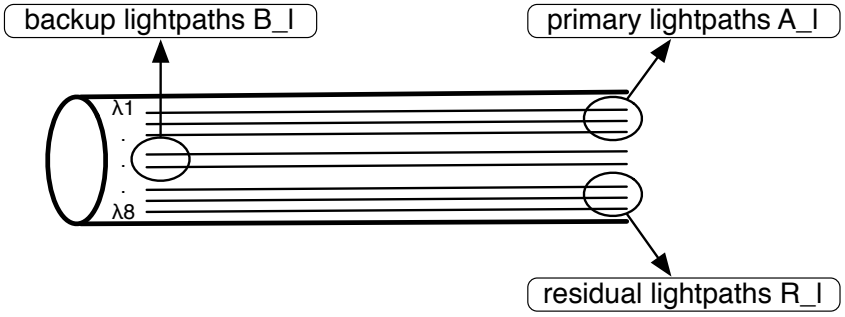

Figure 2.1 – Example fiber link with total capacity of 8 wavelengths

The system model in Figure 2.2 represents visually all the main functions and how these are implemented for each connection request to be assigned a lightpath. In the following sub-sections, we describe each module. Moreover, we provide the model for dual link failures and the metrics used for the performance evaluation.

2.2.1 RWA Problem

The network is represented by a graph $G = (V, L)$ where V is the set of nodes represented by optical cross connects (OXCs) and L is the set of links interconnecting the nodes. Each link is a directed edge from source to destination node and for that reason we assume a pair of such links between each interconnected pair of nodes, in opposite directions. Moreover, each link is assumed to be a single fiber with a capacity of 40 wavelengths. The routing and wavelength assignment problem is solved in two separate steps. Routing for both primary and backup path of each demand is based on Dijkstra's algorithm with

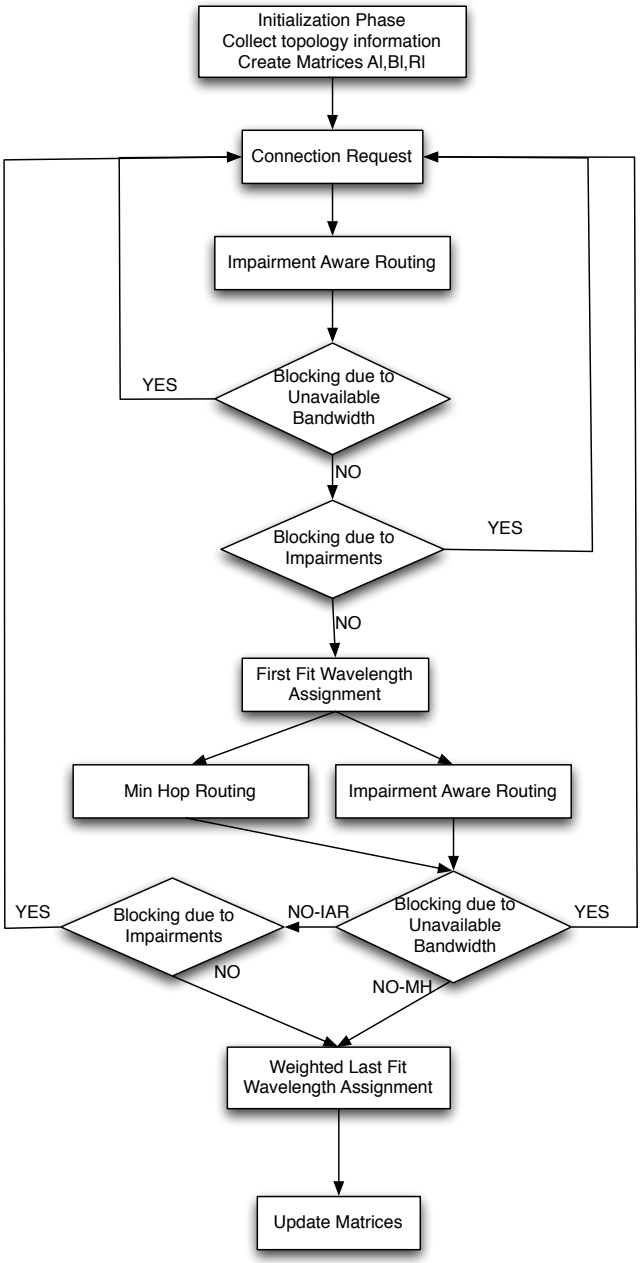

Figure 2.2 – System Model

appropriate link weight assignment: for minimum hop routing, unity weight, whereas for the impairment-aware scheme, the inverse of the analytically computed Q factor [18] of the link. The k-shortest path routing principle is applied in order to increase the

probability to find a feasible path, either in terms of capacity or in terms of a common wavelength across its links. A value of 2 is used for k in this context, thus for each path assignment we compute two candidate paths, applied both for primary and protection paths. The second candidate path is computed after the first is removed from the graph, so the two candidate paths are link-disjoint. The selection of k=2 might seem small for a common graph, but it is considered valid for core network topologies assuming that the candidate paths have no common links, since their connectivity does not usually permit for a much larger number of alternative paths and in the same time the complexity of the algorithm is kept in reasonable levels. After the routing step has identified enough capacity for a lightpath to be established, a wavelength assignment algorithm is applied to select a common -if available- wavelength across each candidate path. Each candidate path computed by the routing algorithm is checked in order of increasing weight for an available common wavelength across all its links. The lightpath is established after the successful completion of both routing and wavelength assignment algorithms, otherwise the connection is blocked.

Physical Layer Impairments Model

The model of physical layer impairments used in this work has been previously presented in [18]. Based on this, a dispersion management scheme and Erbium-Doped Fiber Amplifiers (EDFAs) are used to achieve an acceptable signal quality across the multiple links of a path. The signal is getting distorted in various ways as it propagates through nodes (OXCs and OADMs) and links (fibers employing EDFAs). The degrading phenomena that are taken into account in the present model are: Chromatic Dispersion, Optical Filtering , Amplifier Spontaneous Emission (ASE) noise and crosstalk and WDM non-linearities such as Four-Wave Mixing (FWM), Self-Phase Modulation (SPM) and Cross-Phase Modulation (XPM). A Q-factor metric has been incorporated in order to integrate all the aforementioned types of degradation in a single performance metric. The analytical calculation of the Q-factor is based on the formula presented in Eq. (2.1)

[18].

$$Q_k = \frac{pen_k P}{\sqrt{\sigma^2_{ASE,k} + \sigma^2_{crosstalk,k,k} + \sigma^2_{XPM,k} + \sigma^2_{FWM,k}}} \tag{2.1}$$

where pen_k is the relative eye closure attributed to optical filtering and SPM/GVD and $\sigma^2_{XPM,k}$, $\sigma^2_{FWM,k}$, $\sigma^2_{ASE,k}$ and $\sigma^2_{crosstalk,k}$ are the electrical variances of the degradations induced by XPM, FWM, ASE and crosstalk, respectively.

Primary Path RWA

The primary path of each request is computed according to the IAR scheme by taking into consideration both the available bandwidth and the quality of the paths based on their bit error rate (BER) [18]. The BER of the paths is calculated through the quality factor Q and compared against a predefined BER threshold ($B_{thres} = 10^{-15}$) to decide whether their quality is acceptable. This mechanism enables us to assign link weights that correspond to the physical performance of the links and provide these as input to the routing algorithm, giving this way preference to links that offer higher signal quality. In terms of wavelength assignment the First Fit (FF) algorithm is applied for the primary path.

Backup Path RWA

In the backup path computation phase, the IAR scheme already described is used and compared to the traditional minimum hop routing. The paths computed applying both schemes are compared to the predefined BER threshold and only these with acceptable quality are used; otherwise, blocking due to unacceptable signal quality occurs. Regarding wavelength assignment, we use the Last Fit [19] algorithm that has been proved to maximize the backup path link reuse. However, we have modified the algorithm (Weighted Last Fit) to perform on the possible alternative paths in order of increasing cost. A request is not established unless both a primary and a backup lightpath meeting the signal quality requirements are discovered.

2.2.2 Reinforced Backup Sharing

As already mentioned, our model implements the backup path sharing technique [19] by allowing sharing of the backup link-wavelengths. This way, the backup path of a connection request can use link-wavelengths that are already assigned to backup paths of established requests as long as their primary paths are link-disjoint. In the presence of single link failures, this model provides 100% survivability, as a single link failure will never affect two primary paths that have no common links, requiring the use of a single backup link for the protection of two distinct connections. Figure 2.3 provides a schematic representation of how SBPP works. Paths p_1 and p_2 are two link disjoint primary paths assigned to two different connections. Assuming one wavelength per link being the connection demand volume, the SBPP scheme enables the utilization of a single common wavelength for the two backup paths b_1 and b_2 on link 3-4. If sharing was not permitted, we should have utilized two wavelengths on this link, and in case only one was available, another backup path (probably of higher cost) should have been utlized for one of the two connections' backup path.

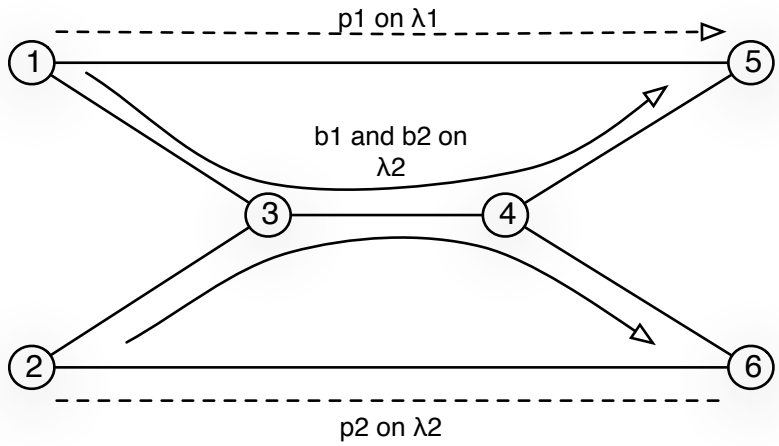

Figure 2.3 – How SBPP works

It is clear that the backup path protection scheme offers a significant gain in terms of

resource utilization efficiency compared to alternative dedicated path protection schemes [28]. In order to further increase this gain, we apply a reinforced sharing mechanism that gives higher preference to links already in use by backup paths during the backup computation phase. This is performed through appropriate assignment of link weights in both impairment-aware and min-hop routing schemes, aiming at higher sharability degrees. Figure 2.4 provides an example of how this works. We assume that primary and backup paths for connection with id 1 from node 1 to node 5 are already assigned as p_1 and b_1 respectively and these are the only active paths present on the topology of the figure. Assuming then another connection request with id 2 from node 2 to node 6, the algorithm works as follows: among the two alternative paths for the primary path of connection 1, p1 is assigned to be the direct link connecting the source and destination nodes; among the two alternative paths for the protection of p_2, represented as choice 1 and choice 2 in the figure and even if choice 2 provides a path with lower cost than the one of choice 1 in the common case, the assignment of lower cost on link 3-4 because of backup path b_1 already using it will result in a lower cost for choice 1. This way, paths that contain links used as protection paths for already established connections are given higher preference during the backup path computation process. Applying this scheme, we aim to further investigate the efficiency of the backup path sharing recovery mechanism and realize possible limits and trade-offs.

2.2.3 Dual Failure Model

Although dual link failures are not as frequent as single failures, they can occur due to different, not uncommon reasons such as shared risk link groups and single failures during link maintenance operations. Considering the huge number of wavelength channels operating in a single fiber link, dual link failures can cause severe loss of connectivity that is projected in serious potential SLA (Service Level Agreement) violation from the operators' side and loss of income. Two link failures that both affect the network at the same time period - in a way that they cannot be treated separately - can occur either concurrently or sequentially overlapping in time. Concurrent failures occur at the same

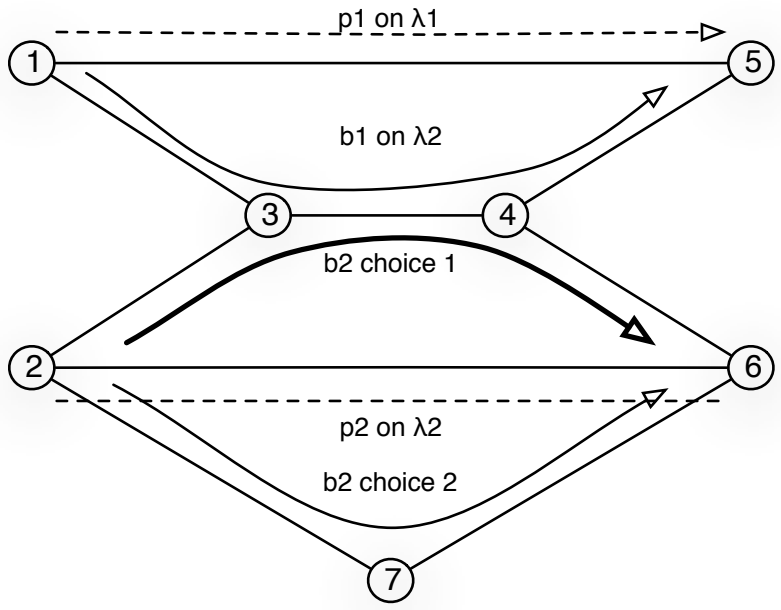

Figure 2.4 – How reinforced sharing works

point in time, thus both find the network in the same state. No further state change occurs between them and the subsequent network state faces the result of both failures. On the other hand, in case of sequential double failures overlapping in time, a failure can arise after a previous failure has already occurred, but before the first failure is repaired. The network may and will probably have changed its state several times between the two failures, but the response to the second failure has to take into consideration the existence of the first failure.

This work focuses on concurrent dual link failures and quantifies their impact on a single failure-resilient network for various routing/sharing scenarios. At each iteration step of the simulation, a link pair is randomly selected and removed from the graph, modeling this way the dual-link failure. After the dual failure realization, a restoration

mechanism takes action and tries to recover the affected connections by computing alternative paths that utilize the spare resources of the network. In the case of sequential overlapping cases, the same metrics would be applicable, but a different implementation should take place that would take into account the state change of the network between the two failures.

2.2.4 Metrics

This work provides a detailed performance evaluation of four different scenarios described below in Section 2.3.1. To ensure the validity of the results the simulation setup includes several repetitions of the same experiment for each connection arrival rate. Random number generators are appropriately initialized to ensure the randomness of each experiment. An initial transient removal mechanism is applied, so that steady state results are ensured, whereas a sufficient number of repetitions is executed so that the 80% confidence interval becomes less than 20% of the sample mean blocking probability.

The total blocking probability in the presence of single link failures is a result of blocking due to unavailable bandwidth on either primary or backup paths and blocking due to physically impaired (compared to the threshold) primary or backup paths.

Utilization is a metric used to illustrate the performance of the different schemes in terms of resource utilization efficiency. Following the notation described in Table 2.1, we subsequently specify two performance metrics used throughout our performance evaluation, namely Sharing Ratio and Utilization.

For each link i, the number of link-wavelengths used is equal to the number of link-wavelengths used for working paths the number of link-wavelengths used for protection paths, both shared and not shared.

$$lw_{used_i} = \sum (lw_{w_i} + lw_{p_i}) \tag{2.2}$$

Table 2.1 – Symbols Notation

Symbol	Description
l_{tot}	Total number of links in the network
W	Total number of wavelengths per link
$l_{w_{tot}}$	Total number of link-wavelengths in the network
l_{w_w}	Link-wavelengths used for primary(working) paths
l_{w_p}	Link-wavelengths used for protection paths
$l_{w_{ps}}$	Shared link-wavelengths used by protection paths
$l_{w_{pns}}$	Not shared link-wavelengths used by protection paths
$l_{w_{used}}$	Total number of link-wavelengths used in the network (working and protection)

Also, for each link i, the total number of protection link-wavelengths is equal to the number of shared link-wavelengths plus the number of not shared link-wavelengths.

$$lw_{p_i} = \sum (lw_{ps_i} + lw_{pns_i}) \tag{2.3}$$

For the entire network, the total number of available link-wavelengths is equal to the total number of links times the number of wavelengths per link:

$$lw_{total} = l_{tot} \times W \tag{2.4}$$

For a whole network instance, utilization is defined as the total number of used link-wavelengths over the total number of link-wavelengths available in the network:

$$U = \frac{\sum_{i=1}^{l_{tot}} lw_{used_i}}{lw_{total}} \tag{2.5}$$

Connection loss rate due to double link failures is used in the context of this work to reflect the extent to which the network is affected by concurrent dual link failures. It is computed as the number of affected connections over the number of active connections in the network, after the removal of the failed link pair. This metric illustrates the vulnerability of the network to dual link failures.

A restoration mechanism is additionally applied in an effort to find an alternative path to reroute connections that were affected by the double failure, using the spare

resources of the network. The fraction of failed connections that are restored is reported as the dual failure restorability. Special attention has to be paid to the way that a dual link failure may affect the established connections. The first and most obvious case is when the link failures affect the primary and the backup path of a single connection. The second case is associated with the fact that backup path sharing is in place. Consider two established connections whose backup paths share some link-wavelengths. If the dual-link failure affects both of these primary paths, the result is that both connections attempt to use the same backup capacity at the same time. One solution could be that we arbitrarily choose one of them to assign the backup capacity and allow the other to be the only connection affected. In this work, we consider both connections failed if such a condition occurs, so that both have to be recovered.

2.3 Results and Discussion

2.3.1 Scenarios

Following the description of the system model and the algorithms used, a detailed presentation of the evaluation scenarios is provided together with the relevant simulation results.

The four scenarios defined for this study are: Sc1 with IAR for primary and backup path computation and reinforced sharing enabled. Sc2 with IAR for both primary and backup paths and conventional sharing. Sc3 with IAR for primary path, min-hop for backup path and reinforced backup sharing enabled. Sc4 with IAR for primary path computation, min-hop for backup and conventional backup sharing. These scenarios are summarized in Table 2.2.

The purpose of these four scenarios is to study the impact of the routing scheme applied for backup path computation on the overall network performance, in combination with the impact of the reinforced sharing mechanism. The resilience of the network in

Table 2.2 – Simulation Scenarios

Scenario	Backup Path Routing Algorithm	Reinforced Backup Sharing
Sc1	IAR	Yes
Sc2	IAR	No
Sc3	min hop	Yes
Sc4	min hop	No

the presence of dual link failures is examined both for networks designed to handle single link failures as well as for networks that incorporate mechanisms addressing dual link failures. More specifically the latter network type is based on shared protection for the case of single failures and activation of a restoration mechanism in the occurrence of dual failures. The two cases are compared and the relevant merits of the two approaches are discussed.

The results presented in this section are generated based on two network topologies, namely the 11-node and 26-link COST 239 [39] and the 16-node and 24-link NSFNET networks, shown in Figure 2.5 and Figure 2.6 [40] respectively.

Particular to the NSFNET topology, regenerators were placed along each link every 600km to avoid unacceptable signal degradation due to physical layer impairments. This is a reasonable assumption to avoid turning entire links unusable due to unacceptably high BER rates.

2.3.2 Blocking Probability

Blocking probability can occur either due to unavailability of bandwidth or due to unacceptably high BER either for the primary or the backup path. Figure 2.7 and Figure 2.8 show all the blocking contributions and their sum (total blocking) for a representative scenario (Sc1) and for both tested networks to demonstrate the above finding. (Notations 'BwPrim'and 'ImpBkp'denote blocking due to bandwidth unavailability on the primary path and blocking due to physical impairments on the backup path, respectively). The contributions not shown in both Figure 2.7 and Figure 2.8 (e.g. blocking due to bandwidth unavailability on the backup path) were found to be zero.

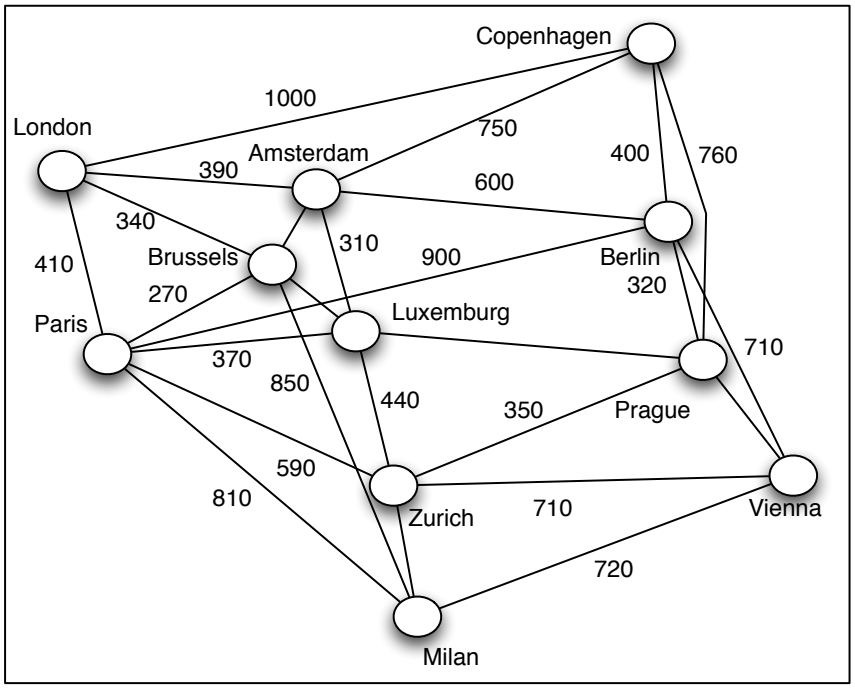

Figure 2.5 – COST 239 Topology

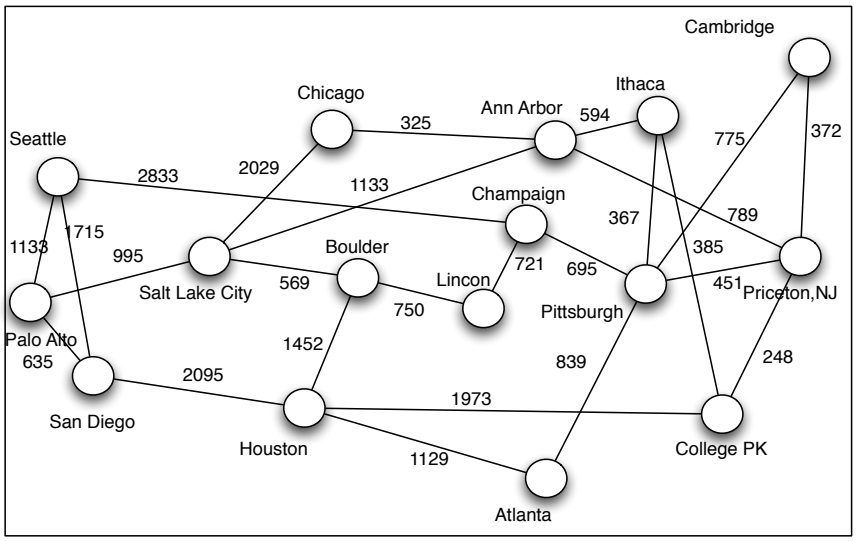

Figure 2.6 – 16-nodes NSFnet Topology

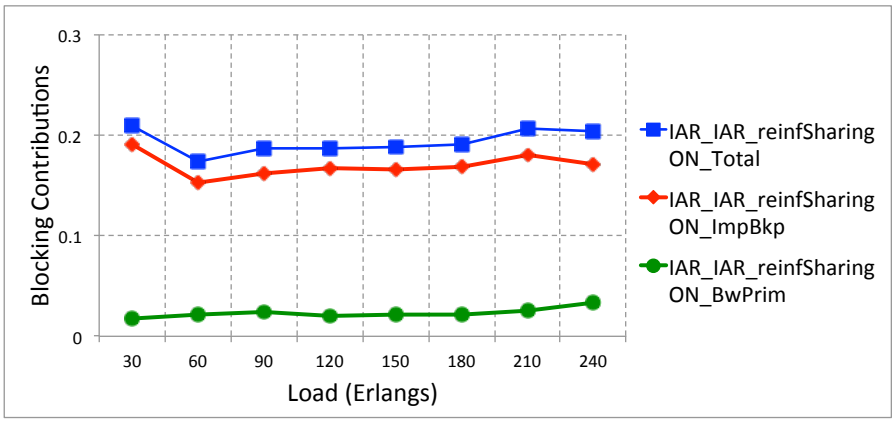

Figure 2.7 – Blocking Contributions - COST 239

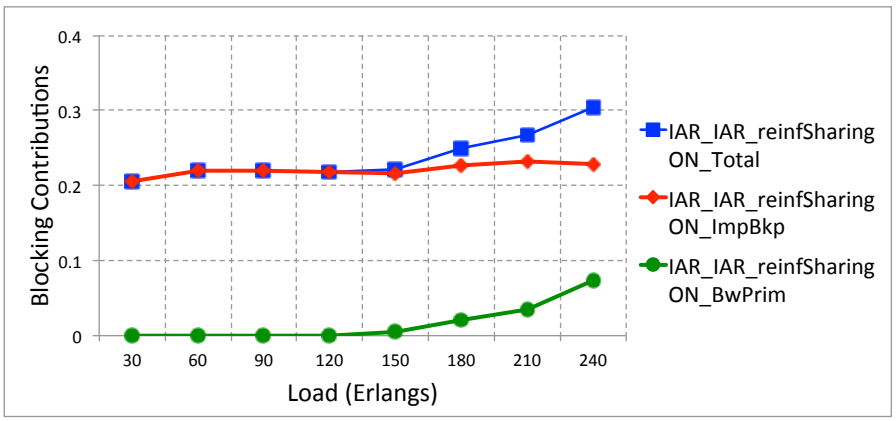

Figure 2.8 – Blocking Contributions - NSFnet

Figure 2.9 and Figure 2.10 illustrate average total blocking probability for a network that is fully resilient in case of single link failures – i.e. without considering the effects of dual failures - for all four scenarios and for both COST 239 and NSFNET topologies respectively. We observe that when IAR was applied for the backup path computation, the total blocking probability for the COST239 topology was much lower compared to the min-hop; for all loading conditions the blocking probability is reduced by more than 50%.

Comparing the scenarios in pairs that follow the same routing principle in the backup path computation phase and aiming at identifying the impact of reinforced backup sharing on the blocking probability of the network (still considering only single link failures),

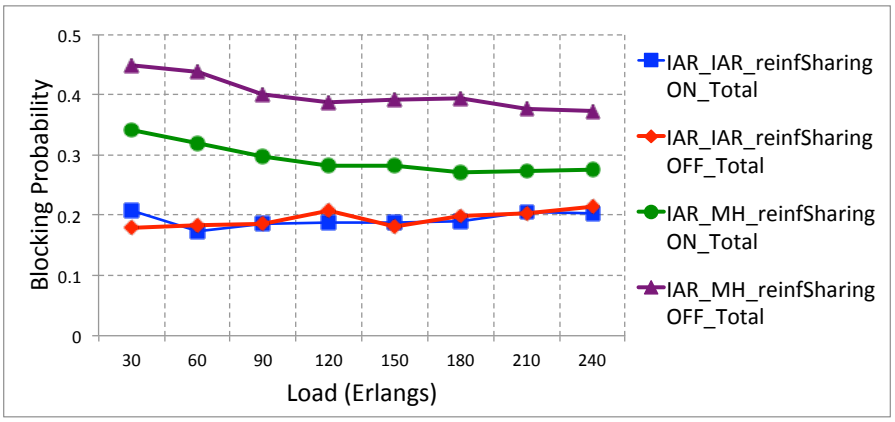

Figure 2.9 – Blocking Probability - COST239

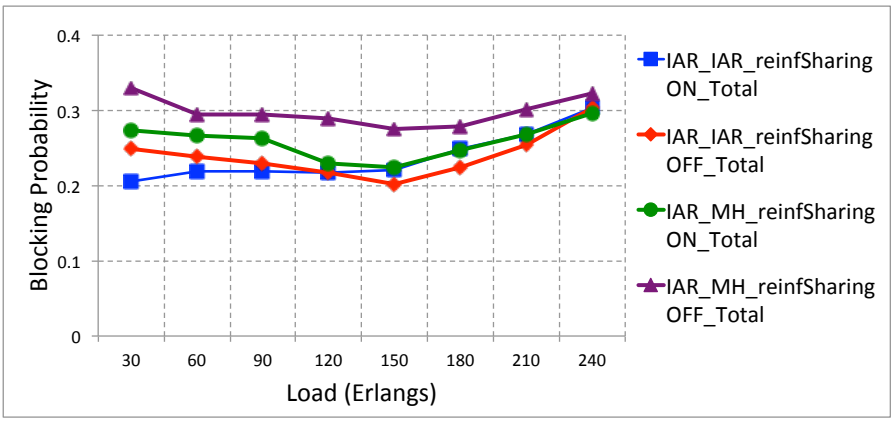

Figure 2.10 – Blocking Probability - NSFnet

we observe the following: Scenarios 1 and 2 that apply IAR in both primary and backup path computations behave in a very similar fashion in terms of blocking when their only difference lies in the application of the reinforced sharing scheme and for both tested networks. This means that the route selection of the backup paths based on physical layer impairments leaves no space on the reinforced sharing scheme to improve the resource utilization. However, the ability of the reinforced sharing scheme to improve network utilization is clearly apparent when the backup path routing scheme is based on the common minimum hop algorithm. Scenarios 3 and 4 result in very different blocking probabilities across the whole load range, with the reinforced sharing scheme demonstrating an important benefit. The result set as presented in this work does not directly

illustrate the resource utilization benefit since the network capacity is considered as given and for all scenarios we try to accommodate the maximum number of connections. If however we demonstrated results with the same blocking, the benefit of the reinforced sharing scheme would be apparent on the resource utilization.

In the NSFNET topology, the IAR also demonstrated reduction in blocking probability, but to a smaller extent, particularly for increased traffic load. This was mainly due to the lower average nodal degree observed in the NSFNET network topology case that implies the existence of a lower number of alternative (link-disjoint) paths for each source-destination pair. The results regarding the efficiency of the reinforced sharing scheme also follow a similar trend to COST 239 topology, again with a smaller difference between scenarios 3 and 4, due to the lack of alternative paths.

2.3.3 Utilization

In addition to blocking, the average usage of resources achieved by each of the four scenarios is studied. Figure 2.11 shows average utilization of the network for all four scenarios and for the COST 239 topology.

Before considering the reinforced sharing mechanism (scenarios 1 and 3), we observe that the min-hop routing in the backup path computation leads to lower bandwidth utilization compared to the IAR scheme. This implies that the IAR scheme utilizes resources more efficiently compared to min-hop routing. This counter-intuitive result can be explained by the fact that the difference in the blocking probability values are attributed to the impaired backup paths as well as bandwidth unavailability.

The main purpose of the evaluation of this metric was aiming at illustrating the trade-offs in the network performance induced by different combinations of routing/sharing schemes, in terms of resource utilization. The reinforced backup path sharing mechanism was applied to investigate the potential gain in terms of resource utilization and sharing. However, as it can be observed (Figure 2.9, Figure 2.10), the consideration of the BER quality during the routing of connections determines the overall network performance and dominates any other effect such as bandwidth availability and resource sharing

capability.

In the case of the scenarios with min-hop backup path routing, the main source of blocking probability is again the high number of impaired backup paths. This is the leading effect, causing different numbers of established connections for the two scenarios and thus higher utilization for the reinforced sharing scheme.

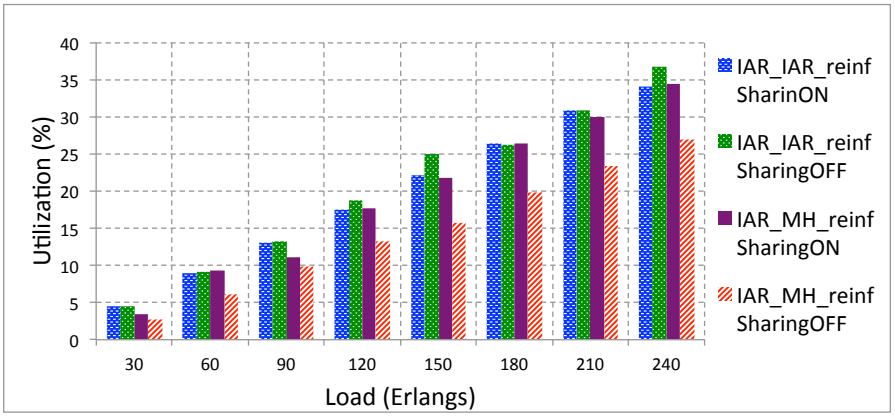

Figure 2.11 – Utilization - COST 239

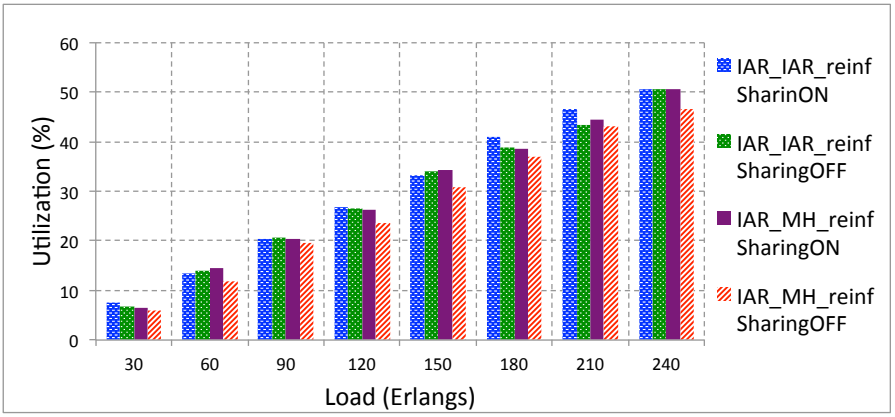

Figure 2.12 – Utilization - NSFnet

In the case where IAR was applied for the computation of the backup paths, the two schemes have similar performance in terms of utilization. This is again due to the dominance of the effect of impaired backup paths. The results corresponding to the

NSFNET topology shown in Figure 2.12 follow a similar trend, without any differentiation in terms of utilization. Again, the utilization results are not considered highly informative in this case, if they are not examined along with the blocking results, since as mentioned above the blocking of connections determines how many connections are established and since this is not the same intentionally for the compared scenarios, the utilization results cannot be directly interpreted.

2.3.4 Dual Failure Connection Loss Rate and Restorability

The results presented so far have clearly identified impairment-aware routing for both primary and backup paths in combination with backup path sharing as the routing scheme that provides the best performance for the scenarios under study and considering single link failures. The investigation is further extended to study the network performance of these scenarios in the presence of dual link failures.

In the context of dual link failure network performance, the most common metric used in the literature is the dual failure restorability. In addition to this, we introduce the connection loss rate, as defined in Section 2.2.4.

Figure 2.13 shows the results for average connection loss rate due to double failures for the COST239 topology. Evidently, all four schemes exhibited similar performance, with connection loss rates in the range of 1 to 1.5%. Particularly, this translates to 1-1.5% of the established and active connections being disrupted by a double link failure, if no additional restoration mechanism is in place. The ability of the network to restore the affected traffic is depicted in Figure 2.14 by the dual failure restorability that yielded results in the range of 60-80%. We observe that the scenarios exhibited performance deviation in high loading conditions, where the existence of the reinforced sharing mechanism resulted in scenarios 2 and 4 worse performance.

In the case of min-hop routing in the backup path, no fair comparison can be made since the two schemes resulted in different blocking probabilities, thus in a different number of established and active connections in the network for the same load. However, in

the case of the IAR scheme in the backup path computation we observe that the reinforced sharing scheme slightly degrades the dual failure restorability performance. The comparison is now possible because of the same blocking of the two scenarios (scenario 1 & scenario 2). We observe that the reinforced sharing scheme degrades the dual failure restorability performance in higher loading conditions. At this point, we have to identify a tradeoff that takes place. Higher resource sharing means that a dual link failure will probably affect more active connections (protected by the same backup links) but on the other hand more capacity will be available in the restoration process. Clearly, the first phenomenon is stronger than the second, resulting in around 10% better dual failure restorability of the conventional sharing scheme.

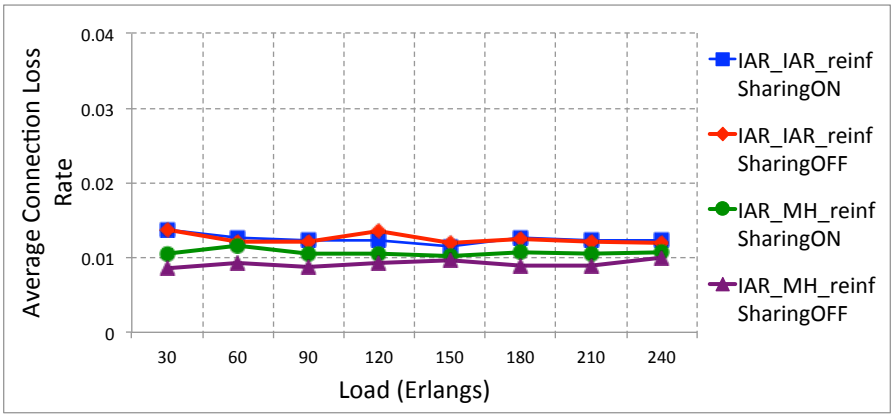

Figure 2.13 – Average Connection Loss Rate - COST 239

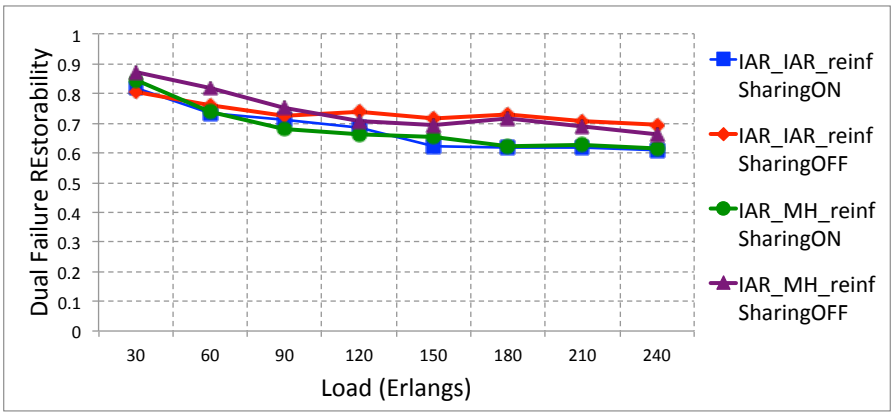

Figure 2.14 – Dual Failure Restorability - COST 239

Figure 2.15 and Figure 2.16 show the same set of results for the NSFNET topology, with average connection loss rate in the range of 3-3.5% and restorability in the range of 20-60%, respectively. The difference in the dual failure restorability compared to the COST239 network relates to the lower nodal degree of the NSFNET that limits the number of possible alternative paths between each source-destination node pair. Both the impairment-aware and the min-hop routing schemes demonstrate similar performance in terms of connection loss rate due to dual failures, whereas the dual failure restorability results are also of the same levels. Schemes 1 and 2 performed better in high loading conditions in these terms for NSFNET, as was the case in COST239.

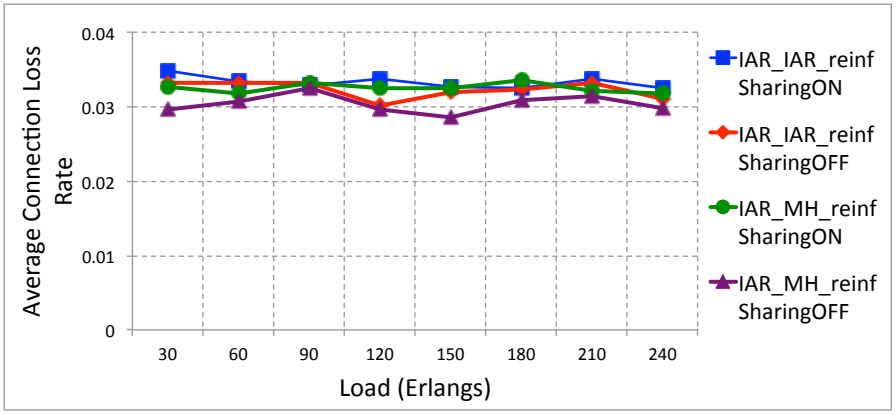

Figure 2.15 – Average Connection Loss Rate - NSFnet

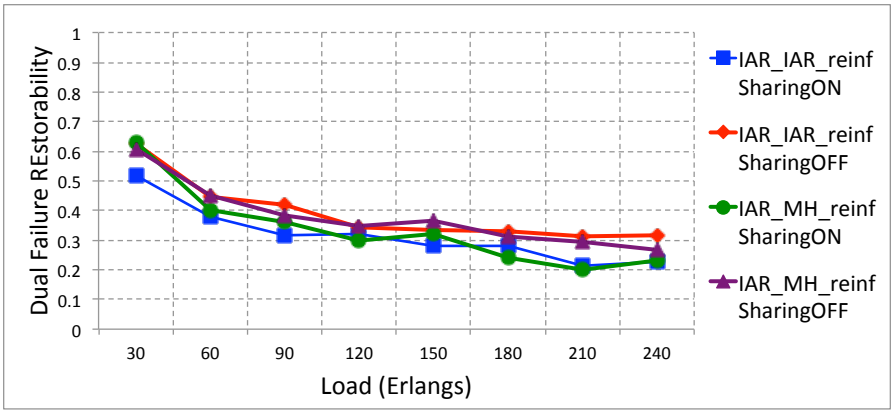

Figure 2.16 – Dual Failure Restorability - NSFnet

The study presented in this work has led into two main findings. The first associates

network performance in terms of blocking and capacity utilization with the physical performance of the network links and the second evaluates the network performance of the different combinations (routing/sharing) in the presence of dual link failures.

An overall evaluation of the results leads to the conclusion that the consideration of impairments in the routing algorithm for both the primary and the backup path provides a reduced network blocking probability and thus an improved overall solution in the presence of single and dual link failures.

2.4 Conclusions

The study presented in this chapter has led to two main findings. The first associates network performance in terms of blocking and capacity utilization with the physical performance of the network links and the second evaluates the network performance of the different combinations (routing/sharing) in the presence of dual-link failures. When the quality of the formed paths is taken into account in the process of routing, the physical performance of the selected paths is shown to be better, compared to the traditional routing schemes that consider only bandwidth availability. Thus the impairment-aware routing scheme, although it does not always identify the shortest path, it provides better performance under single and dual-link failures. The finding is driven by the almost invariant blocking probability of all schemes for a wide loading range. The main contributor to this blocking is the dominant effect of physical impairments in the backup paths for the different simulated scenarios. The results show a relatively constant response of the network (both COST 239 and NSFNET) to incoming requests for the whole loading range. Regarding the dual-link failures, although the absolute values of the average connection loss rate imply high tolerance of single-link failure-resilient networks in the presence of dual failures, spare capacity placement or restoration schemes against dual failures have to be seriously considered, since the obtained results are sufficient to violate the required 99.999% availability of carrier-grade networks. An overall evaluation

of the results leads to the conclusion that the consideration of impairments in the routing algorithm for both primary and backup paths provides a reduced network blocking probability and thus an improved overall solution in the presence of single- and dual-link failures.

Chapter 3

The Impact of Optical Wavelength Conversion on the Energy Efficiency of Resilient WDM Optical Networks

As the on-going growth of bandwidth demands has been leading the expansion of the Internet in size and complexity, 4% of the primary energy worldwide has been identified as the percentage of ICT power consumption worldwide contributions [49]. This growth brings along increased energy consumption by the network, alerting the interest of the community in energy efficient networking and altering the main focus of energy-efficiency from access to core networks [50]. In this context, engineering communication networks in a power-aware manner seems to be instrumental for a more energy-efficient Internet. Although optical networking is an energy efficient technology itself, the need for power-awareness in core optical network design and protocol implementation in addition to the requirement for energy efficient system design has been demonstrated in the literature [51] along with the benefits of equipment selective equipment switching off [21], [52].

Fault tolerance is a major aspect in optical WDM networks whose fiber link capacity reaches 40 and 100 wavelengths to date and link and/or node failures can disrupt a huge

part of established communications if not mitigated. For this reason and as presented in Chapter 2 and the relevant literature, different protection mechanisms that either protectively or dynamically allocate dedicated (1+1,1:1) or shared (Shared Backup Path Protection-SBPP) excess network capacity have been developed to provide resilience against failures. The various developed mechanisms provide different trade-offs between the level of protection offered and the amount of excess capacity needed. It is obvious that this extra capacity has to have some impact on the additional equipment needed to support it and thus the power consumption of the network infrastructure. This is not described by a linear relationship, since for example the same optical amplifier can support different dimensions of fibers in terms of the number of wavelengths, while being fully operating and consuming the same power. However, excess capacity may result in more fibers per link which means that more amplifiers will be needed for the connection of two nodes, whereas the optical node size may require significant extensions to support the additional fibers.

In this context, the work of this chapter aims to address the impact of protection in WDM optical mesh networks on the total power consumption and quantify the savings offered by the use of all-optical wavelength conversion technology instead of the traditional optical-electrical-optical converters. By dimensioning the network through minimum-cost optimization problems for unprotected, dedicated and shared path protection schemes, we evaluate the impact of the use of optoelectronic and all-optical wavelength converters on the total network power consumption. Relevant power consumption results have been produced assuming 10Gb/s and 40Gb/s per wavelength channel data rates. These results indicate that the use of all-optical wavelength conversion significantly assists in decreasing the overall network power consumption and identify the SBPP scheme as the most energy–efficient survivability solution.

3.1 Related Work

In [49], the authors present the results of a survey regarding the main contributors to the power consumption in Information and Communication Technology (ICT), identifying mainly the power consumption of the relevant equipment during use and the power consumption for manufacturing this equipment. All contributions are estimated to account for about 4% of the primary energy consumption globally. An estimation for ICT equipment consumption is also presented, forecasting a percentage of more than 14% in the year 2020. Finally, research challenges are identified regarding hardware power optimization as well as software-related issues including intelligent power management and energy efficient network design and operation. In [51], after identifying system design, network design and protocols as the main areas for power-awareness, a generic router power consumption model is presented. Based on this, the authors provide network design models that aim to obtain results regarding the associated power needs of the network based on specific topologies and demands. Using these results, they propose re-routing rules that enable the network-wide minimization of the power consumption and provide an important space for power savings. In [52], Chiaraviglio et al. present optimization problems that aim to minimize the power consumption of wide-area networks by exploiting over-provisioning of resources and switching off of equipment. By selectively switching off nodes and links, they demonstrate power savings in the order of 50% and 30% respectively. The authors of [20] formulate and solve appropriate optimization problems that aim to minimize the energy consumption of an IP over WDM network by exploiting lightpath bypass in the optical domain. Significant savings in the order of 25% to 45% are demonstrated due to lightpath bypass, whereas IP routers are illustrated to account for 90% of the total network power consumed.

3.2 Problem Formulation

3.2.1 Capacity Placement Network Design

The network is modeled as a graph $G = (N, L)$ comprising a set of nodes $N(n = 1, ..., N)$ and a set of links $L(l = 1, .., L)$ interconnecting the nodes. The nodes are Optical Cross-Connects (OCXs) based on 3D MEMS [53] switching technology, as illustrated in Figure 3.1, employing OEO regenerators (transponders) at every output port supporting also 100% wavelength conversion capability. The set of nodes are interconnected by a set of unidirectional links L, whose length is also given as an input.

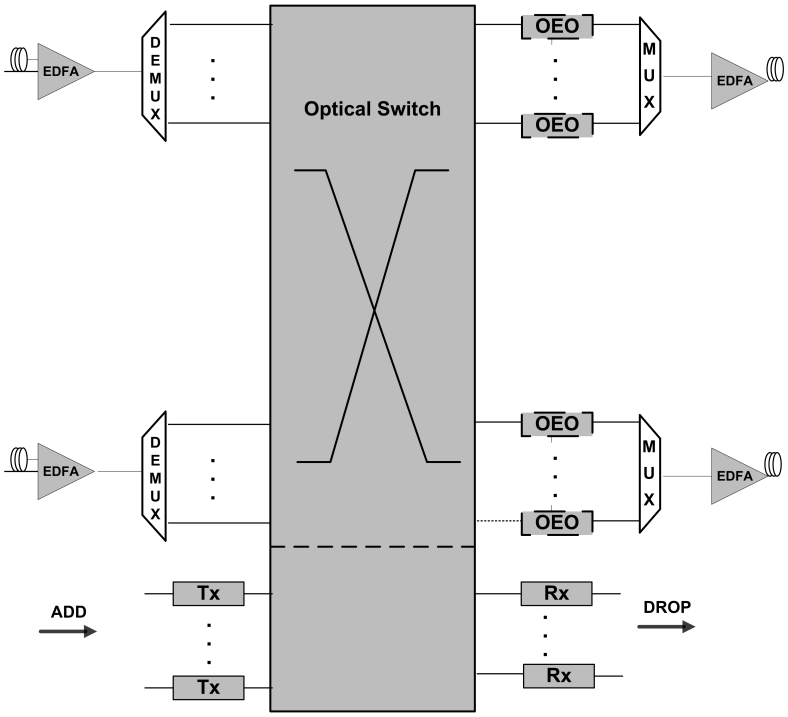

Figure 3.1 – Optical Cross Connect Architecture

A traffic matrix (TM) represents the traffic demands between the nodes of the network. Each value of the TM is an integer number, since it is assumed that all demands are integer multiples of wavelengths. The node pair is in the form (Origin O, Destination D) and each request is directional. In the context of the work presented in this chapter, a "link" is considered to be the set of fibers installed between two interconnected nodes

whereas a fiber-link is a specific fiber on that link. The capacity of each link is not given, since this is the result of the capacity placement problem. Thus, the problem solution defines the number of fibers needed to be installed on each link. However, the maximum fiber capacity is given as an input, as the maximum number of wavelengths W per fiber link. The cost values of the different network equipment are also given. This involves mainly the cost of the fiber links and the OXC equipment.

Working on a regime of 70 to 280 average number of lightpath requests per source-destination pair that follow a uniform distribution, we form a set of traffic volume instances, aiming at addressing worst-case growing traffic between the network nodes. We formulate and solve three problems dealing with : a) an Unprotected network where each demand is assigned a lightpath from the source to the destination node, b) a Dedicated Protection scheme where each demand is assigned a primary and a dedicated link-disjoint backup lightpath and c) a Shared Backup Path Protection scheme where each demand is assigned a primary path and a link-disjoint backup path whose capacity can be shared with other backup lightpaths, assuming that the corresponding primary paths are link-disjoint. Length-based shortest path routing is employed to populate the candidate path lists for both primary and backup paths. Each wavelength demand is assumed to be of 10Gb/s granularity and two scenarios are considered, one supporting 10Gb/s and one supporting 40Gb/s per wavelength channel data rates. For both protection schemes, the primary path is assigned to be the first shortest path, while for protection, k=2-shortest path routing is used to populate the set of candidate backup paths.

The geographical location of the nodes and the trenching cost for the fiber links are not considered, assuming that the physical links interconnecting the nodes are already in place. We further assume that sharing of add/drop ports between the primary and backup lighpath of each demand is taken into account. The output of the network dimensioning process provides the optimal number of fibers per link and wavelengths per fiber needed to serve the input traffic matrix at minimum cost. From this output, the optimal dimensions of optical switches required are also calculated. In the subsections following

Table 3.1 – Problem Variables

Variable	Domain	Description
CAP_l	Z_+	Number of wavelengths established on link l
w_l^k	0,1	=1 if link l is used on the primary path of demand k
f_l	Z_+	Number of fibers installed on link l

Table 3.2 – Problem Parameters

Parameter	Description
$\omega(n)$	set of links incident to node n
$\omega^+(n)$	set of links leaving node n
$\omega^-(n)$	set of links arriving at node n
$O(d)$	Origin node of demand d
$D(d)$	Destination node of demand d
W	Fiber link maximum capacity (wavelengths)
h_d	Demand d volume

we present the exact Integer Linear Programming (ILP) formulations for unprotected, dedicated and shared path protected networks.

Objective Function

The objective of all problems in this chapter is to minimize the total number of fibers installed in the network to accommodate the traffic requests. This is illustrated in the objective function as depicted in Eq. (3.1). Table 3.1 and Table 3.2 illustrate the basic set of variables and parameters respectively used in the formulation of all subsequent problems. We assume single wavelength volume (h_d=1) for each demand d. Additional variables are used when needed.

$$minimize \sum_{l \in L} f_l \tag{3.1}$$

Unprotected Scheme

The formulation for the unprotected case is presented here in the form of a node-link capacity placement problem. The problem consists of three basic constraints sets, the flow conservation constraints (Eq. (3.2)) ensure that the flow required to accommodate the connection demand is conserved across the nodes of the lightpath. The constraints presented in Eq. (3.3) assign the required number of wavelengths on each link interconnecting a node pair. Thus the total number of fibers needed to be installed between two interconnected nodes is depicted on the variable CAP_l and this drives the correct assignment of the variable depicting the number of fibers to be installed between these nodes, that is f_l in Eq. (3.4).

Flow Conservation Constraints

$$\sum_{l \in \omega^+(n)} w_l^k - \sum_{l \in \omega^-(n)} w_l^k = \begin{cases} -1, \text{if } n = O(d) \\ +1, \text{if } n = D(d) \\ 0, \text{otherwise} \end{cases} \quad n = 1, ..., N, d = 1, ..., D \tag{3.2}$$

Link Capacity Constraints - Wavelengths

$$\sum_{d \in D} w_l^k h_d \leq CAP_l, l = 1, ..., L \tag{3.3}$$

Link Capacity Constraints - Fibers

$$f_l \geq \frac{CAP_l}{W} \tag{3.4}$$

Dedicated Path Protection

The formulation for dedicated 1:1 path protection follows the same principles as the one presented above for the unprotected case with a set of variables and constraints added. Table 3.3 presents the additional variables and the additional and/or modified constraints. This way, constraints Eq. (3.5) ensure flow conservation also for the backup capacity of the dedicated path protection scheme. Link capacity for the backup part of

Table 3.3 – Additional Problem Variables - Dedicated Path Protection

Variable	Domain	Description
$BCAP_l$	Z_+	Number of wavelength established on link l as backup capacity
b_l^k	0,1	=1 if link l is used on the backup path of demand k

each link is also assigned according to Eq. (3.6) constraints, whereas Eq. (3.7) ensures the mutual exclusion of primary and protection paths in the sense that the primary and the backup path that serve a single connection have to be link-disjoint, as a basic principle of the recovery scheme aiming 100% protection against single link failures. The number of fibers to be installed on each link is now driven from both working and protection capacity requirements as Eq. (3.8) depicts.

Flow Conservation Constraints - Backup Path

$$\sum_{l\in\omega^+(n)} b_l^k - \sum_{l\in\omega^-(n)} b_l^k = \begin{cases} -1, \text{if } n = O(d) \\ +1, \text{if } n = D(d) \\ 0, \text{otherwise} \end{cases} \quad n = 1, ..., N, d = 1, ..., D \tag{3.5}$$

Link Capacity Constraints - Wavelengths - Backup Path

$$\sum_{d\in D} b_l^k h_d \leq BCAP_l, l = 1, ..., L \tag{3.6}$$

Mutual Exclusion Between Primary and Backup Flow Variables

$$w_l^k + b_l^k \leq 1 \tag{3.7}$$

Link Capacity Constraints - Fibers - Primary and Backup Path

$$f_l = \frac{CAP_l + BCAP_l}{W} \tag{3.8}$$

Table 3.4 – Additional Problem Variables - Shared Path Protection

Variable	Domain	Description
$v^b_{d,p}$	0,1	additional binary variable used for linearization of backup capacity constraints

Shared Path Protection

Following the exact same node-link formulation principles of the dedicated path protection presented above, the formulation of the shared path protection scheme needs the insertion of one additional variable, as presented in Table 3.4. The reason for that is that the original formulation requires the multiplication of two binary variables in the constraint set that assigns wavelengths for backup capacity resulting in non-linear constraints. The additional variable $v^k_{d,p}$ along with the additional constraints in Eq. (3.10) enable us to transform the problem to linear and solve it with traditional ILP solvers.

Link Capacity - Backup Flows -Wavelength

$$\sum_{d \in D} v^b_{d,p} h_d \leq BCAP_l, l = 1, ..., L \tag{3.9}$$

Linearization

$$w^k_l + b'^k_l \leq 1 + v^b_{d,p}, d = 1, ..., D, l = 1, ..., L, l' = 1, ..., L, l \neq l' \tag{3.10}$$

3.2.2 Power Consumption Model

After deriving the optimal network dimensions for the considered input traffic, we compute the total power consumption for each node n as a function of four factors: (a) the switch fabric (P^n_{SF}), the transponders for (b) transmission (P^n_{Transm}) and (c) wavelength conversion (P^n_{Conv}) and (d) the optical amplifiers (P^n_{Ampl}). The power consumption values of the power-dissipating network elements presented in Table 3.5 are obtained from

related literature [21] and vendors' data sheets. Eq. (3.11) - Eq. (3.14) illustrate the network power consumption computation for each network node n, whereas Eq. (3.15) presents the respective calculation for each fiber link l connecting two nodes. For each fiber, a span length of 80 km (span_length) is assumed. f_{in}^{n} denotes the number of fibers arriving at node n and f_{out}^{n} the number of fibers leaving node n.

Switching Fabric

$$P_{SF}^{n} = ports_{total} \times P_{port\_pair} = (ports_{th} + ports_{a/d}) \times P_{port\_pair} \tag{3.11}$$

Transmission

$$P_{Transm}^{n} = ports_{a/d} \times P_{transpX}, \text{where } X\text{=10G or 40G} \tag{3.12}$$

Wavelength Conversion

$$P_{Conv}^{n} = ports_{th} \times P_{Y}, \text{ where } Y\text{=transp}X \text{ } (X\text{=10G or 40G}) \text{ or SOA} \tag{3.13}$$

Optical Amplification

$$P_{Ampl}^{n} = (f_{in}^{n} + f_{out}^{n}) \times P_{edfa} \tag{3.14}$$

Fiber Link

$$P_{l} = \lfloor \frac{length(l)}{span\_length} \rfloor \times P_{edfa} \tag{3.15}$$

The total network power consumption $P_{Network}$ is then calculated as in Eq. (3.16).

$$P_{Network} = \sum_{n} P_{n} + \sum_{l} P_{l} \tag{3.16}$$

$P_{n} = P_{SF}^{n} + P_{Transm}^{n} + P_{Conv}^{n} + P_{Ampl}^{n}$ is the total power consumption for node n and P_{l} is the power consumption for link l.

Table 3.5 – Power Consumption of Active WDM Network Components

Symbol	Description	Power Consumption
P_port_pair	Input/Output port pair of 3D MEMS switch fabric	0.107 mW
$P\_transp10G$	OEO Line-side WDM Transponder (10G - 80km)	6 W
$P\_transp40G$	OEO Line-side WDM Transponder (40G - 80km)	18 W
P_SOA	SOA-based all-optical wavelength converter	3.5 W
P_edfa	EDFA Optical Amplifier	13 W

3.3 Results and Discussion

Without imposing any bound on the number of fibers to be installed per link, we compute the optimal network dimensions for each scheme and at a post-processing stage we calculate the total network power consumption based on these dimensions. As a reference topology the Pan-European COST 239 (Figure 2.5) network topology has been used, comprising 11 nodes and 52 unidirectional links, whereas the fibre link capacity is 40 wavelengths.

The results in Figure 3.2 and Figure 3.3 clearly show that the use of optical wavelength conversion provides significant energy savings in the network for all different resilience schemes considered and the two data rates examined i.e. 10Gb/s and 40Gb/s. Moreover, the SBPP scheme outperforms the 1:1 path protection scheme for both 10Gb/s and 40 Gb/s as expected. At 40 Gb/s the performance of the SBPP scheme is close to the unprotected case, thus providing a solution that is both energy-efficient and resilient against single link failures. In more detail, for 10Gb/s the energy savings remain always above 25% for the unprotected case and this percentage increases with the volume of traffic that is supported by the network. When examining the SBPP and 1:1 protection schemes, this percentage always exceeds 28% and again increases further for higher traffic volumes. At 40Gb/s the effect follows a similar trend, but it is more profound as in this case the savings always exceed 33% and reach a maximum of 58% for networks supporting high volumes of traffic and offering 1:1 protection.

To offer some detailed insight on the various equipment contributions to the total network power consumption, Figure 3.4 - Figure 3.9 provide a relevant breakdown for all

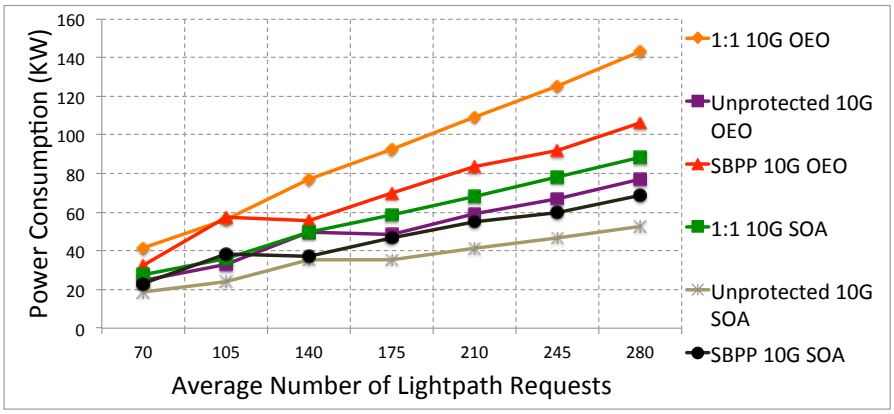

Figure 3.2 – Total Power Consumption - 10G - OEO vs SOA

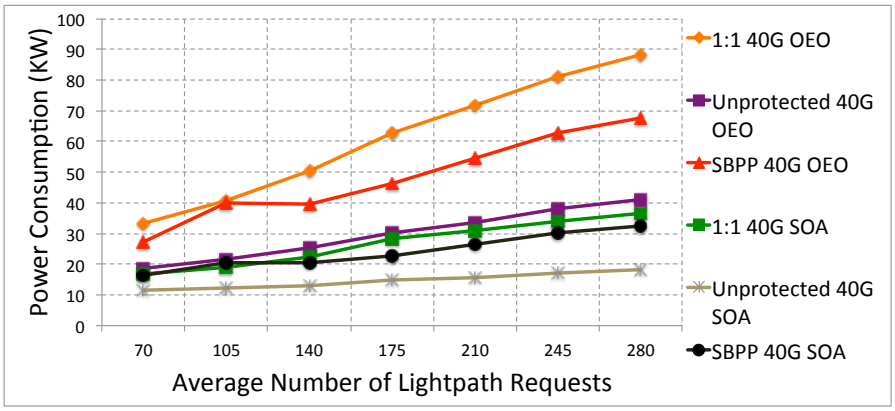

Figure 3.3 – Total Power Consumption - 40G - OEO vs SOA

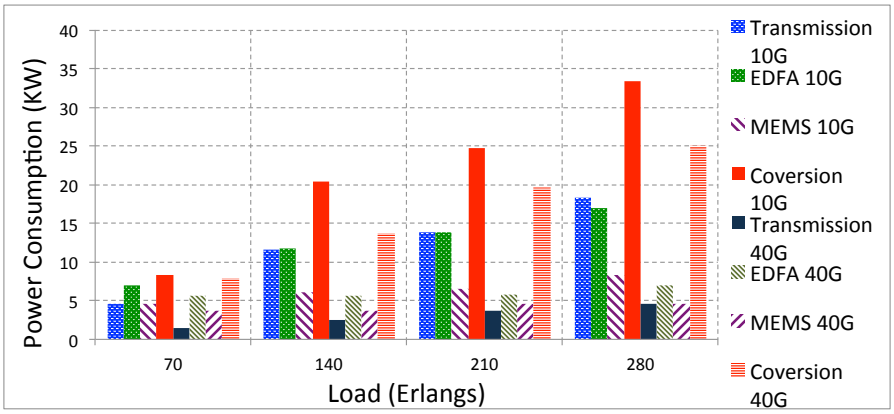

Figure 3.4 – Power Consumption - Breakdown by Element - Unprotected OEO Conversion

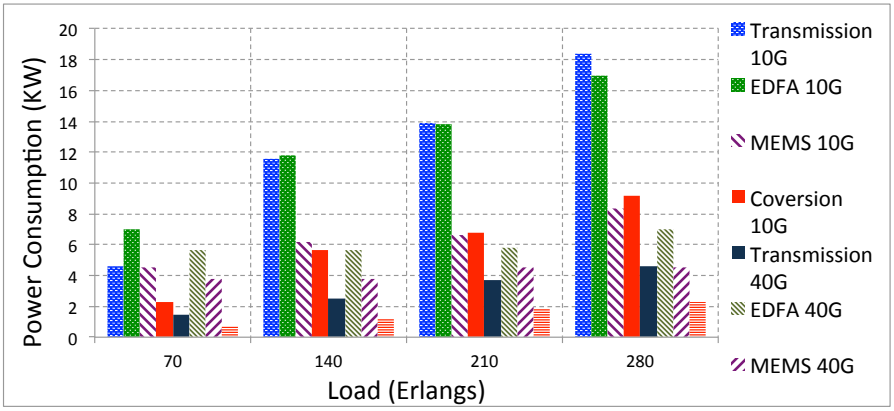

Figure 3.5 – Power Consumption - Breakdown by Element - Unprotected SOA Conversion

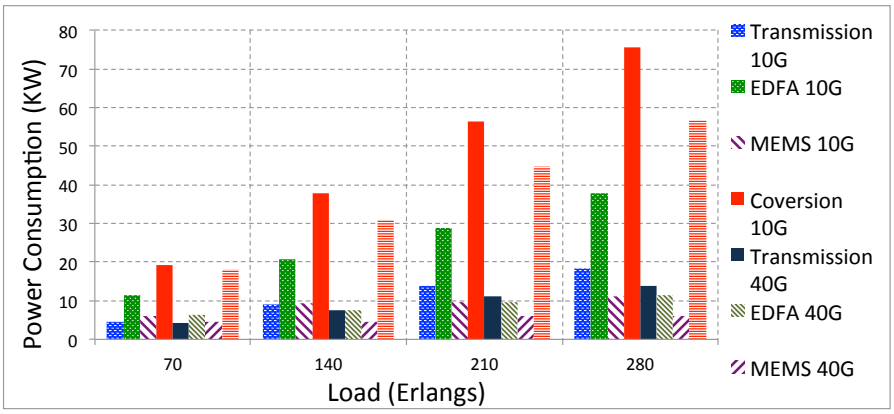

Figure 3.6 – Power Consumption - Breakdown by Element - 1:1 OEO Conversion

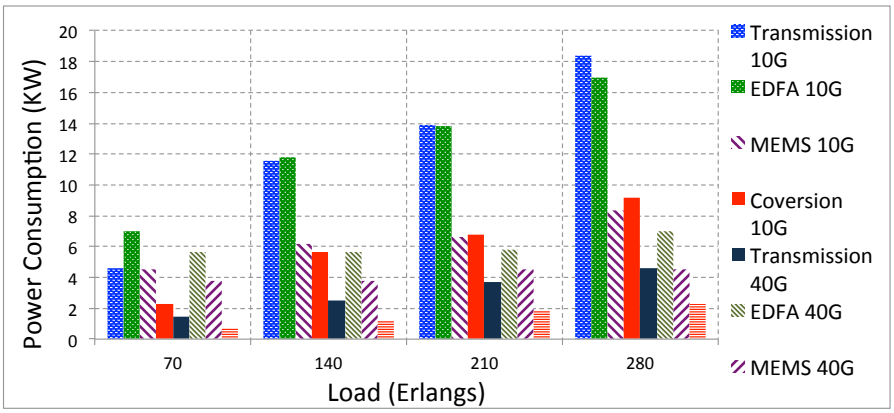

Figure 3.7 – Power Consumption - Breakdown by Element - 1:1 SOA Conversion

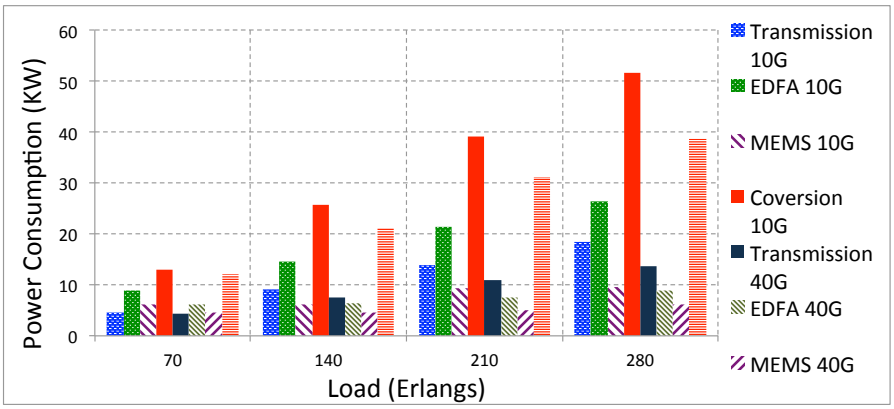

Figure 3.8 – Power Consumption - Breakdown by Element - SBPP OEO Conversion

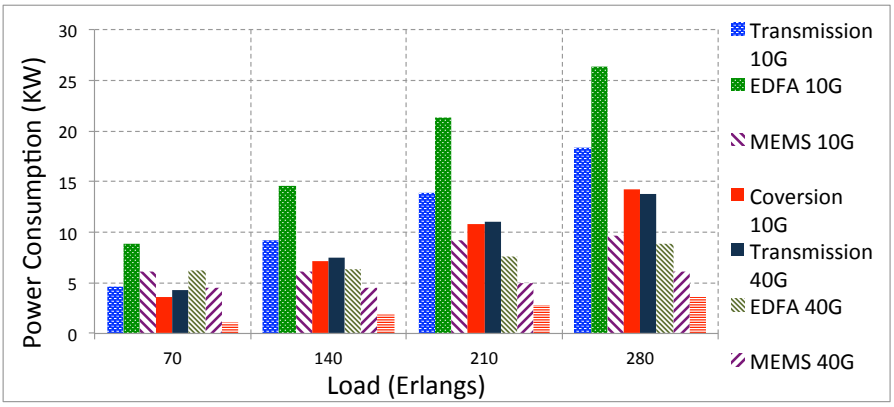

Figure 3.9 – Power Consumption - Breakdown by Element - SBPP SOA Conversion

three cases of unportected, dedicated and shared protection, as well as for both OEO and SOA-based wavelength conversion. The results clearly indicate that the use of optical wavelength conversion facilitates the overall energy savings observed. These figures also give a map of the various optical components energy footprint, from a network perspective, identifying that depending on the traffic volumes supported by the network, optical components that have a significant contribution to the overall energy consumption are optical amplifiers and transmitter/receiver pairs. Finally, they illustrate the change in the percentage of the total power consumption that the different network elements are assigned to, identifying a raise for the optical amplifiers and the transmission equipment

after the adoption of all-optical wavelength conversion.

3.4 Conclusions

The work presented in this chapter evaluated the impact of all-optical wavelength conversion technology on the overall power consumption of resilient WDM optical networks. After considering the cases of no protection, 1:1 dedicated path protection and shared backup path protection schemes, we demonstrated that the shared path protection scheme achieves increased efficiency in the utilization of network resources. In combination with the energy-efficient technology of all-optical wavelength conversion, we illustrated how this scheme provides resilience to single link failures at the expense of minimal increase in the total power consumption of the network compared to the unprotected case.

Chapter 4

Converged Optical Network and Data Center Infrastructure Planning

Cloud-based services are increasingly deployed taking advantage of the continuous advancements of data centers (DCs), while large-scale service providers like Amazon and Google increasingly deploy geographically dispersed DCs [54] to satisfy the requirements of the offered services. These services include storage, processing, e-mail, web-services and gaming, whereas private enterprise DCs are also used for data-intensive tasks like Web page indexing and large data-sets analysis [55], [56], [57]. On demand self-service, location independence, rapid elasticity, reliability and disaster recovery are some key-requirements [58] that the cloud infrastructure needs to satisfy. It is true to say that cloud infrastructures have emerged as an evolution step from computing grid infrastructures, adopting some of the main technologies and approaches used to serve these requirements. These include the most promising optical networking architectures in terms of technology, as well as advanced routing, virtualization, control [57] and joint consideration of network and computing resources [59].

In this context, the concept of virtualization applied to an infrastructure comprising DCs interconnected through a WDM optical network supporting cloud services, can offer

performance advantages [60] and facilitate sharing of physical resources. This enables the introduction of new business models [61] that suit well the nature and characteristics of the Future Internet and enable new exploitation opportunities for the underlying physical infrastructures (PIs). In this environment service providers are able to establish their own virtual infrastructures (VIs) over the underlying physical infrastructure.

In the converged infrastructure described above there are two types of resources: a) network resources including fiber links and nodes, and b) DC resources comprising storage, processing cores and memory. The infrastructure planning process is usually formulated as an optimization problem with common objectives dealing with resource utilization. Efficient resource sharing, minimum resource allocation and load balancing are variants of this objective class. Previous work, [56], [59] mainly addressing grid computing solutions has already identified the joint consideration of these two types of resources as the most effective way towards efficient utilization of the infrastructure. However in cloud computing, fault management and load-balancing usually require all available DCs to be active and the optimization related to resource utilization to take place internally within the DCs.

As Information and Communication Technology (ICT) is estimated to be responsible for more than 4% of the primary energy worldwide, a percentage expected to significantly increase by 2020 [62], a lot of attention has been recently paid on the energy efficiency of such converged infrastructures [56], [61]. Optical networking is an energy-efficient technology that can be further optimized with regards to energy consumption through power-aware network design and protocol implementations [63]. However, the operation of DC resources requires very high levels of power and their conventional operating window is commonly not optimized for energy efficiency. Allocating Information Technology (IT) processing jobs in an energy-aware manner through a relatively low energy-consuming optical network and switching-off unused IT resources can offer significant energy savings.

This chapter focuses on the design of virtual infrastructures (VIs) over a physical infrastructure (PI) taking into consideration jointly the network and DC resources. The

novelty of this work lies in a more realistic assumption that there is no global knowledge of the requests for all the VIs, thus performing the planning of each VI in sequence according to the arrival order of the VI requests over the underlying PI that is already supporting previously established VIs. Through the design process, both the topology and required virtual resources are identified and mapped to the physical resources and the associated operating parameters. We compare two objectives, one minimizing the joint power consumption of network and DC resources (MJP) and one minimizing the network resources used (MNR). The goal of this comparison is to identify suitable design objectives, tradeoffs and trends for realistic VI request scenarios and a variety of traffic loading conditions. Moreover, we study the impact of the design objectives on the resulting virtual topologies and their performance under dynamic traffic. Our results demonstrate that although the MJP objective achieves lower power consumption compared to the MNR as expected, the benefit decreases as the number of established VIs and the volume of demands supported increases. The performance comparison of the different planned VIs shows that the gain in the power consumption offered by the MNR objective introduces a penalty in the blocking performance.

4.1 Background and Related Work

4.1.1 Energy Efficiency and Optical Networks

In [64] the authors present a breakdown of energy consumption in cloud computing spanning from network-related switching and transmission to data center-related data processing and storage. Transport-related energy consumption is identified as an important contributor to the total cloud energy consumption. In [56] the authors provide a comprehensive study of the major approaches to achieve energy efficiency in telecommunications networks by describing the main elements in each of the three network domains (core, metro and access) and their energy consumption. They provide a breakdown of the main components that consume energy (core routers, optical cross connects

and optical amplifiers) and identify optical switching as the most energy-efficient solution. Regarding core networks, they present the main approaches to achieve energy efficiency, such as selective turning off of network equipment, energy-efficient network design and green routing. Finally, the authors focus on data centers and respective applications where power control of high-speed network intra-connection and optimal load-distribution schemes across data centers are referenced as proposed solutions for intra- and inter-data center communications. In [65] the authors demonstrate the efficiency of joint network and IT consideration in terms of power consumption over the physical infrastructure and demonstrate a benefit of 3-55%. This benefit depends on the ability of a data center to switch on/off servers and is compared to the case where only network power consumption is minimized. In [61], a model describing the concept of planning virtual infrastructures over of a converged network and IT infrastructure is presented and an energy-aware VI planning problem is formulated and compared to approach that allocates the IT services to the server located closest to the source node, providing savings of the order of 30% for a single VI establishment. Moreover, energy aware offline service provisioning for the VWP case on top of the planned VI is presented and compared to three other approaches, achieving minimum overall power consumption. Finally, in [66] the authors propose multiple VI energy-aware planning, assuming global knowledge of the VI requests and provide the optimal solution for the establishment of all VIs concurrently.

4.1.2 Optical Networking in Grid and Cloud Computing

In [67] the authors demonstrate a joint computational and network resource scheduling system based on an experimental network (TONICE) that supports distributed scientific applications, proposing two scheduling approaches. The first proposed scheme is overlay scheduling, where network and computational resources are considered separately and computational resources are the basis of scheduling, whereas the second scheme, that is integrated scheduling jointly considers computational and network resources aiming at better performance and resource utilization. The experimental results demonstrate that

the integrated approach out-performs the overlay solution in terms of average schedule length, computational resource utilization and optical network resource requirements. In the grid networking context where a set of geographically distributed computing nodes are interconnected through a WDM network and logical partitions (tasks) of application instances (jobs) have to be transferred and executed within a time deadline, the authors of [59] deal with the problems of task scheduling and lightpath establishment jointly, formulating and solving a joint optimization problem referred to as the "Task Scheduling and Lightpath Establishment"problem. They propose two algorithms, one that minimizes the job completion time and one that minimizes the cost subject to a deadline constraint. Their results demonstrate better performance in the former case against traditional list scheduling. In [57] the authors start with a presentation of the applications that drive the evolution from grid to cloud computing. These mainly involve scientific applications (scientific and data-centric computing), business applications (transactional systems, multimedia, data mining) and consumer applications (gaming, interactive TV, augmented reality). They continue with the main requirements for optical grids/cloud arising from the aforementioned applications, identifying mainly on-demand setup, response time, latency, scalability and elasticity among others. Moreover, they present the migration from grid to cloud, based on the need to distribute not only scientific but also consumer applications and provide the main cloud paradigms (Software as a Service, Platform as a Service and Infrastructure as a Service). Finally, they present the technologies that enable the realization of the services fulfilling the identified requirements in terms of optical technology, routing, virtualization and control and management.

4.2 System Model

4.2.1 Network Model

The network is modeled as a graph, comprising a set of nodes N interconnected by a set of unidirectional links L. The nodes are Optical Cross-Connects (OXCs) based on a photonic switching matrix that is realized by 3D Micro-Electro-Mechanical Systems

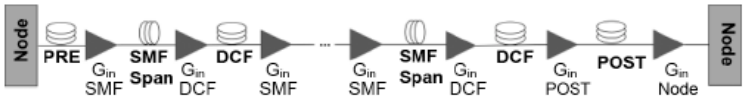

Figure 4.1 – Fiber Link Model

(MEMS) as in Figure 3.1.

Each node supports M input and M output fibers, each employing a maximum number of wavelengths W. Apart from the passive elements, namely the Multiplexers (MUX) and de-multiplexers (DEMUX), Figure 3.1 illustrates the active elements of the OXC: the switch matrix, one Erbium-Doped Fiber Amplifier (EDFA) per input/output fiber port, one Optical-Electrical-Optical (OEO) transponder per output wavelength port and one transmitter (Tx) - receiver (Rx) pair per lightpath. The OEO transponders are used to support the wavelength conversion functionality in the case of the VWP network. The number of through (express) ports is calculated as the number of input fibers M times the fiber wavelength capacity W. It is further assumed that the add/drop capability of the node is 50% of the through traffic. No OEO converters are included in the WP network architecture.

Figure 4.1 illustrates the link architecture [17] employed for the interconnection of the OXCs. It is modeled as a sequence of alternating Single Mode Fiber (SMF) and Dispersion Compensation Fiber (DCF) spans, to address fiber dispersion effects including Pre- and Post-compensation DCF spans at the beginning and the end of each link respectively. To compensate for the insertion loss of the fiber spans optical amplifiers based on Erbium Doped Fiber Amplifier (EDFA) technology are allocated at the end of each transmission span.

4.2.2 Data Center Model

The main building block of the data center model is based on [68]. It is a full rack implementation of a hardware platform that is used in real DCs and its main characteristics are summarized in Table 4.1. For the data center throughput, we have assumed it to be

equal to four times the 75 GB/s per rack uncompressed I/O bandwidth reported in [69].

Table 4.1 – Data Center Building Block Characteristics

Resource	Type	Capacity
CPU	30 servers/360 cores	2.93 GHz Intel Xeon 6-core processor
Storage	-	40 TB
Tdc	-	2.4 Tbps
Memory	1333 MHz	2.9 TB

4.2.3 Power Consumption Models

Network

The total network power consumption is determined by the power consumption of OXCs and fiber links. Figure 3.1 illustrates the OXC architecture and its power-dissipating elements with grey color. The node power consumption (P_{OXC}) depends on four factors: (a) the switch fabric (P_{SF}), the transponders for (b) transmission (P_{Transm}) and (c) wavelength conversion (P_{Conv}) and (d) the optical amplifiers (P_{Ampl}). Eq. (4.1) to Eq. (4.4) demonstrate the computation of these power consumption values, whereas Table 4.2 provides a short description and typical power consumption values for the required equipment [70],[71].

$$P_{SF} = ports_{total} \times P_{portpair} = (ports_{th} + ports_{a/d}) \times P_{portpair} \tag{4.1}$$

$$P_{Transm} = ports_{a/d} \times P_{Tx/Rx} \tag{4.2}$$

$$P_{Conv} = ports_{th} \times P_{Transponder} \tag{4.3}$$

$$P_{Ampl} = (f_{in} + f_{out}) \times P_{Edfa} \tag{4.4}$$

In our case, we assume a symmetric switch MxM whereas M is not identical across the OXCs, but computed after the planning process based on the traffic volume supported.

Table 4.2 – Network Equipment Power Consumption Figures

Symbol	Description	Power (Watts)
P_portpair	Input/Output port pair of the switch fabric	0.10710^{-3}
P_transponder	O/E/O Line-side WDM Transponder	6
P_Tx/Rx	E/O,O/E Transmitter or Receiver	3.5
P_edfa	EDFA	13

The only power consuming elements included within the optical network links (Figure 4.1) are the optical amplifiers installed per span. The amplifier span length is assumed to be 80km. Thus the power consumption of a fiber link is length-dependent and is calculated as depicted in Eq. (4.5):

$$P_l = \lfloor \frac{length(l)}{span} \rfloor \times P_{Edfa} \tag{4.5}$$

The total network power consumption of the physical infrastructure is computed by Eq. (4.6):

$$P_{Net} = \sum_{n \in N} P_{OCX} + \sum_{l \in L} P_l \tag{4.6}$$

Data Center

The power consumption of the data center is based on typical power consumption values taken from [68]and on the simple linear model illustrated in Eq. (4.7).

$$P_{DC} = P_{idle}(P_{busy} - P_{idle}) \times u_s \tag{4.7}$$

P_{idle} is the power consumption of the DC in idle state, that is when no server is used.

Its value is considered to be the 50% of the DC power consumption under full utilization, P_{busy}. In our model, this value is assumed to be the maximum power consumption value from [68], as shown in Table 4.2. Finally, u_s represents the utilization of a data center s that is defined as the sum of lightpath requests currently served by the data center over the DC throughput expressed in wavelengths. Eq. (4.8) illustrates the computation of the DC utilization.

$$u_s = \frac{\text{number of lightpaths arriving at s}}{T_{DC}} \tag{4.8}$$

Table 4.3 summarizes the power consumption values used for each DC, assuming the same configuration and capacity for all data centers employed in the converged infrastructure.

Table 4.3 – Data Center Power Consumption Figures

Symbol	Description	Value (KW)
P_busy	Power Consumption under full utilization	17.5
P_idle	Power Consumption under idle state	8.75

4.3 Problem Formulation

4.3.1 Virtual Infrastructure Planning

The virtual infrastructure planning problem is formulated as an integer linear program (ILP) and is based on the well-studied Routing and Wavelength Assignment (RWA) [24] problem. As also defined in the introduction section the term PI refers to the Physical Infrastructure, that is the set of DCs and WDM nodes and links that interconnect them. The term VI refers to the Virtual Infrastructure that is a slice of the PI comprising a set of DC and network resources in terms of DC throughput and wavelengths respectively. In this work we concentrate on optimal planning of virtual infrastructures with respect to specific objectives, while further virtualization implementation details are not taken

into account. More specifically we provide two sets of formulations using path and flow variables, one for the virtual wavelength path and one for the wavelength path case.

Virtual Wavelength Path

Let $G = (N, L)$ be the directed graph that represents the network topology, where N is the set of network nodes (OXCs) and L is the set of directed links interconnecting the nodes. Let also S be the subset of N that represents the nodes where data centers are attached and D be the set of requests for each virtual infrastructure. Multiple sets, each one corresponding to a VI request, are treated sequentially in the order of arrival of the VI requests without any prior knowledge of the VI requests. We solve one optimization problem for each virtual infrastructure that needs to be formed and established over the converged physical infrastructure. The solution of each problem updates the network and DC capacity and gives a new instance of the infrastructure as an input to the next problem. The solution of the problem provides an optimal mapping of the virtual infrastructure to the physical infrastructure, both in terms of network topology and infrastructure resources, according to the objective function.

The planning problem considered in this work is treated as a capacitated problem. We start with a given DC capacity offered over the entire infrastructure and a relative network capacity that is sufficient to accommodate the traffic requests aiming to utilize the DC resources. The definition of the demand volume needs to include the amount of resources requested, both in terms of network and data center. The characteristics that enable the definition of an exact relationship between these two types of resources are the DC throughput and the wavelength bit rate. Eq. (4.9) gives the number of wavelengths that are needed to satisfy the DC throughput.

$$k = \frac{T_{DC}}{R_W} \tag{4.9}$$

Each demand d of the set D is described by the source node and by the constant h_d that is the demand volume assuming wavelength level granularity. For each virtual

infrastructure problem, the demand sources are randomly selected between all the network nodes apart from the ones that are directly connected to DCs. Each problem is solved for a set of total lightpath requests that span from 30 to 150, whereas for each such value the reported results correspond to averages over a number of 50 repetitions.

We refer to physical topology links (or physical links) by the index l ($l = 1, ..., L$) and to virtual topology links (or virtual links) by the index e ($e = 1, ..., E$). The physical link set L is straightforward to obtain since it represents all the directed links of the graph G. The virtual link set E is assumed to be the set of links in a full mesh graph with the same set of nodes N. Moreover, network nodes are indexed by n over the set N ($n = 1, ..., N$) and data centers are indexed by s $s = 1, ..., S$. Finally, physical and virtual candidate path lists are indexed by q ($q = 1, ..., Q$) and p ($p = 1, ..., P_d$) respectively.

Apart from the demand volume h_D, we use two more constants in the problem formulation. As already mentioned, the destination is not part of the demand description, since all data centers with available capacity are candidate destinations for every request. Thus for every demand and every candidate destination, a candidate path list is generated based on k-shortest path routing. We refer to this as the virtual candidate path list and use parameter δ_{edp} in Eq. (4.10) to represent the use of candidate virtual path p in the realization of virtual link e for demand d.

$$\delta_{edp} = \begin{cases} 1, \text{if candidate virtual path } p \text{ of demand } d \\ \text{is used to realize virtual link } e \\ 0, \text{otherwise} \end{cases} \tag{4.10}$$

$$\gamma_{leq} = \begin{cases} 1, \text{if physical link } l \text{ is part of candidate} \\ \text{physical path } q \text{ of the virtual link } e \\ 0, \text{otherwise} \end{cases} \tag{4.11}$$

Since we need to map the virtual to the physical network resources, we use one more constant to illustrate the realization of virtual link by the corresponding physical path q, illustrated in Eq. (4.11).

For both candidate path lists, we use a length-based Yen's k-shortest path algorithm [72] and retrieve the two shortest paths from every source to all candidate destinations. Table 4.4 summarizes the variables whose optimal values form the solution of the ILP and represent the resource assignment that accommodates the requested traffic and provides the mapping between physical and virtual resources.

Table 4.4 – Problem Variables

Variable	Domain	Description
x_{dsp}	Z_+	Flow realizing demand d towards DC s on candidate virtual path p (number of lightpaths)
z_{eq}	Z_+	Flow realizing virtual link e on candidate physical path q
w_e	Z_+	Number of wavelengths utilized on virtual link e
y_l	Z_+	Number of wavelengths utilized on physical link l
z_{eq}	Z_+	$\begin{cases} 1, \text{if DC } s \text{ is used} \\ 0, \text{otherwise} \end{cases}$

Two different optimization problems are formulated by using two objective functions, namely "NetRes" in Eq. (4.12) and "JointPower" in Eq. (4.13). As the names indicate, the objectives minimize the total number of wavelengths (representing the network resources) used and the total power consumption of both network and DC resources respectively. Based on the problem variables and the power consumption models, the two objectives are:

$$NetRes = \sum_{l} w_l \times length(l) \tag{4.12}$$

$$JointPower = \sum_{s} \left[(1 - f_s)P_{idle} + (P_{busy} - P_{idle})\, u_s\right] + \sum_{n} \sum_{l \in \omega^-(n)} \left[(1 - \alpha_l)\, P_{link} + \alpha_l P_{link} \left(\frac{w_l}{\sum_l \Lambda_l} + 1\right)\right] \tag{4.13}$$

where

$$P_{link} = \sum_{n} \sum_{l \in \omega^-(n)} \frac{1}{W_l} \left(\lfloor \frac{length(l)}{span} \rfloor + 2\right) P_{Edfa} \tag{4.14}$$

and

$$f_s = \begin{cases} 1, \text{if DC } s \text{ is already used by another VI} \\ 0, \text{otherwise} \end{cases} \tag{4.15}$$

and

$$a_l = \begin{cases} 1, \text{if link } l \text{ is already used by another VI} \\ 0, \text{otherwise} \end{cases} \tag{4.16}$$

The indices $\omega^+(n)$ and $\omega^-(n)$ represent the outgoing and incoming links of node n respectively. Λ_l is the number of already established wavelengths on link l by previous VIs. As indicated by Eq. (4.13), the power consumption of the optical links is calculated as follows: when a VI is the first to utilize a link, the total power consumption of the link is assigned to this VI. When more than one VIs use the same link, each VI is assigned proportionally a power consumption level reflecting the utilization of the corresponding link resources (wavelengths).

The constraints that follow complete the ILP formulation and ensure that network and DC resources are optimally assigned, following flow conservation and capacity rules.

The constraint in Eq. (4.17) ensures that all demands d are served:

$$\sum_{s}\sum_{p} x_{dsp} = h_d, d = 1, ..., D \tag{4.17}$$

The next three constraints in Eq. (4.18))-(Eq. (4.20) deal with capacity bounds and require that the virtual and physical links capacities w_e and y_l respectively, are enough to accommodate all lightpaths noting that these two variables are upper bounded by the fiber capacity; based on commercially available WDM products, the latter is assumed to be 96 wavelengths. Moreover, they provide the mapping between physical and virtual links. The mapping is performed through constraints in Eq. (4.18) and Eq. (4.19) that map the virtual links capacities to the physical links capacities through the flow variable.

$$\sum_{d}\sum_{s}\sum_{p} \delta_{edp} x_{dsp} \leq w_e, e = 1, ..., E \tag{4.18}$$

$$\sum_{q} z_{eq} \leq w_e, e = 1, ..., E \tag{4.19}$$

$$\sum_{e}\sum_{q} \gamma_{leq} z_{eq} \leq y_l, l = 1, ..., L \tag{4.20}$$

Finally, two more constraints ensure the correct assignment of the binary variable f_s that represents whether a data center is used or not.

$$f_s \leq u_s, s = 1, ..., S \tag{4.21}$$

$$f_s T_{DC} \geq u_s, s = 1, ..., S \tag{4.22}$$

The utilization of the DC u_s is defined as:

$$u_s = \sum_{d}\sum_{p} x_{dsp}, s = 1, ..., S \tag{4.23}$$

Wavelength Path

The path formulation for the WP case follows the same principles and notation with the VWP case described above with the addition of the index c that represents the distinct wavelengths of a fiber link. The integer flow variable of the problem that indicates the number of lightpaths using the virtual path p to support demand d that will be serviced by the corresponding DC s is now updated to x_{dspc} and indicates the assignment of a specific wavelength c across the path. Accordingly, the flow variable z_{eq} is updated to z_{eqc}. The rest of the variables, indices and constants remain the same.

The constraints of the problem are updated to ensure the correct assignment of capacity across the paths for the establishment of all the demands: Eq. (4.24) ensures the establishment of all the demands, whereas constraints in Eq. (4.25))-(Eq. (4.27) assign to virtual and physical links the required capacity to accommodate the traffic flows.

$$\sum_{s}\sum_{p}\sum_{c} x_{dspc} = h_d, d = 1, ..., D \tag{4.24}$$

$$\sum_{d}\sum_{s}\sum_{p}\sum_{c} \delta_{edp} x_{dspc} \leq w_e, e = 1, ..., E \tag{4.25}$$

$$\sum_{q}\sum_{c} z_{eqc} \leq w_e, e = 1, ..., E \tag{4.26}$$

$$\sum_{e}\sum_{q}\sum_{c} \gamma_{leq} z_{eqc} \leq y_l, l = 1, ..., L \tag{4.27}$$

Eq. (4.28) ensures that each distinct wavelength on a physical path q realizing a virtual link e is assigned only to one flow.

$$\sum_{e}\sum_{q}\sum_{c} \gamma_{leq} z_{eqc} \leq y_l, l = 1, ..., L \tag{4.28}$$

Eq. (4.21) and Eq. (4.22) of the VWP formulation remain the same. Finally, the

data center utilization in Eq. (4.29) is expressed as the sum of lightpaths arriving at DC s:

$$\sum_{e}\sum_{q}\sum_{c}\gamma_{leq}z_{eqc} \leq y_l, l = 1, ..., L \tag{4.29}$$

The optimization objectives remain as defined for the VWP case.

The problem formulations used for the planning of the infrastructure without considering the formation of VIs are referred to as PI planning problems and are subsets of the formulation presented above. Therefore, they do not include variables and constraints related to the virtual links and paths.

4.3.2 Online Traffic Provisioning

The virtual infrastructures planned through the integer programs based on the two objective functions presented previously are further evaluated under an online traffic provisioning scenario. The simulations are based on a custom Matlab tool [73] and traffic is modeled as a Poisson arrival process with exponential service time distribution. Requests are randomly generated from all sources, apart from the nodes directly connected to data centers. The traffic load spans from 30 to 80 Erlangs.

The goal of these simulations is to evaluate the performance of the different VIs in terms of request blocking. The request destination is a decision based on the closest-DC scheme, where for every request we choose as destination the DC that can be reached through a length-based shortest path with available capacity. The granularity of the requests is one wavelength. The performance evaluation takes place for virtual infrastructures that correspond to both VWP and WP planning problems solutions and both optimization objectives. For the WP case, the First Fit wavelength assignment algorithm is used.

4.4 Results and Discussion

To evaluate the different VI planning approaches described above, we assume the integrated network and DC infrastructure presented in Figure 4.2. The reference network topology used is based on the COST 239 pan-European network [39] interconnecting four data centers. One fiber link per direction ensures bi-directional communication with fiber capacity of 96 wavelengths. The data center throughput in terms of wavelengths is 240 wavelengths. The mapping of the data center and network resources requires that the four data centers' total capacity is 240*4=960 lightpaths and our traffic demands reach the level of 150 lightpath requests per VI, thus 600 lightpaths for the case where all 4 VIs are present. We assume four virtual infrastructure request sets, each one supporting a total number of lightpath requests ranging from 30 to 150. Moreover, we assume that only active (utilized) network and data center resources consume power, whereas the respective granularity is the fiber link and the DC as a whole.

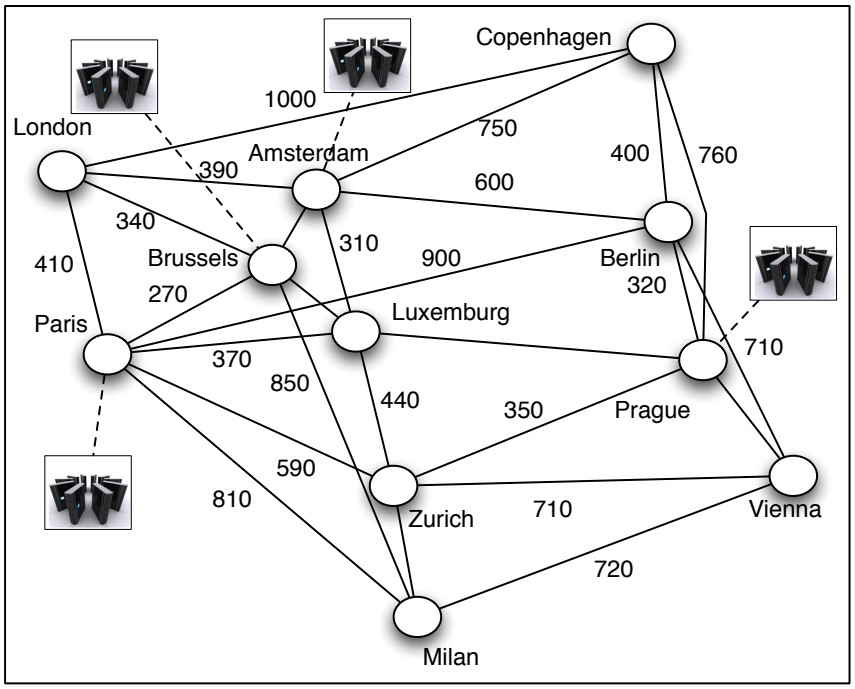

Figure 4.2 – COST 239 Topology with 4 Data Centers

To address the statistical uncertainty associated with the input traffic, we report confidence interval (CI) limits of the mean network utilization for a confidence level of 95%. Each value corresponds to a specific VI and lightpath load combination and is computed as the lower and upper (±) CI limit over the mean network utilization for 50 repetitions. For each repetition, we generate a set of lightpath requests that sum up to the corresponding loading value and the source nodes are randomly selected according to a uniform distribution. Table 4.5 provides the corresponding values for the case of wavelength path formulation and the MNR objective. All other formulations result in similar values that always remain in the range of 3%-15%.

Table 4.5 – CI Limits of Mean Power Consumption for 95% Confidence Level

VI / Lightpaths	30	60	90	120
1	0.1026	0.0282	0.0129	0.0147
2	0.0335	0.0279	0.0129	0.0148
3	0.0335	0.0280	0.0128	0.0145
4	0.0334	0.0278	0.0128	0.0143

4.4.1 Power Consumption

The planning problems presented in this work model two different objective functions (MJP and MNR) under two network technologies (VWP and WP). This study aims at investigating: a) the impact of the VI model and request model on the power consumption of the infrastructure and b) the impact of the objective function to the power consumption when VIs are sequentially planned for both VWP and WP networks. These issues are illustrated in Figure 4.3-Figure 4.13 and are analyzed below:

a) Figure 4.3 and Figure 4.4 demonstrate two result sets for the VWP and WP case respectively and concentrate on the total power consumption corresponding to the planned VIs for each objective and for all four VIs and the breakdown of power consumption on DC and network parts. We first observe that the MJP objective achieves

lower power consumption values across lower traffic volumes, as expected, whereas a change of this trend is observed for high numbers of VIs and load values.

Focusing on the VWP case, it is observed (Figure 4.3) that the MJP objective achieves significantly lower power consumption for 120 and 240 ligthpaths (corresponding to 30 and 60 lightpaths per VI), almost similar values with the MNR objective for 360 and 480 lightpaths, and finally higher power consumption for 600 lightpaths. The breakdown of network and DC power consumption provides a more detailed understanding on how network and DC resources need to be powered up based on the output of the two objectives and explains why the MJP objective does not lead to the optimal power consumption across all load values and VI numbers. In this context, there are two main observations: 1) The MJP demonstrates almost constant network power consumption across all load values (small variations are observed due to statistical error), that is also much higher than the respective network power consumption when applying the MNR objective. This is verified in Figure 4.5 that demonstrates the average lightpath length for the two objectives.

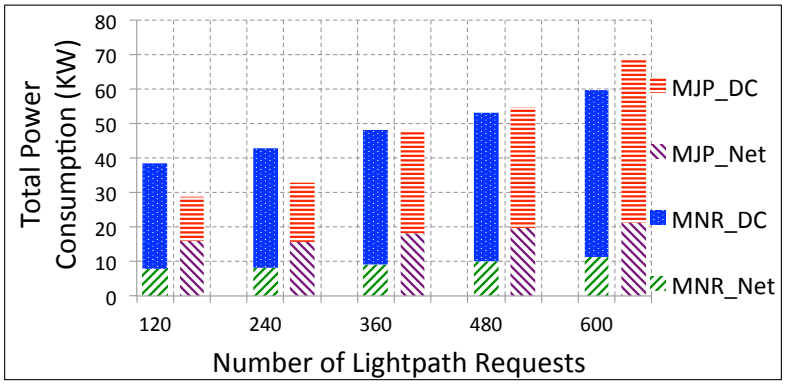

Figure 4.3 – Power Consumption Contributions - VWP

2) At the same time, the DC power consumption of the MJP objective is much lower than that of MNR for lower load values and grows to the same level for higher load values. This is clearly verified in Figure 4.6, where the number of powered up DCs versus the number of requests is illustrated. We observe that the MJP objective causes powering up of DCs only when the already powered up DCs are not sufficient

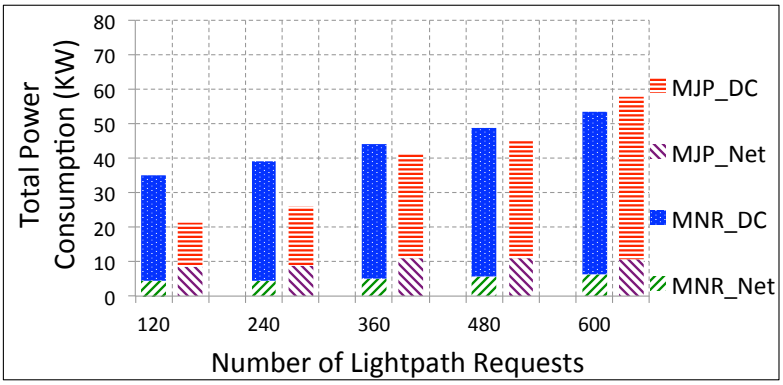

Figure 4.4 – Power Consumption Contributions - WP

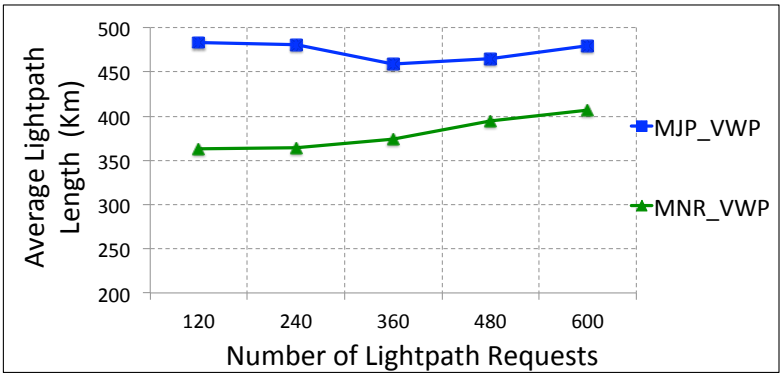

Figure 4.5 – Average Lightpath Length - VWP

to accommodate the new requests. Since the DCs are the highest energy consuming elements of the infrastructure the decision to power up DCs plays an important role in the optimization process and leads to higher average lightpath length.

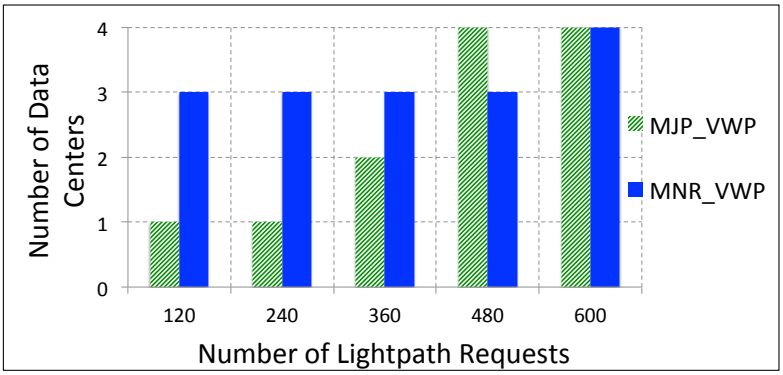

Figure 4.6 – Number of Data Centers Powered Up

These results have demonstrated that the MJP objective gradually causes powering up of DCs and introduces longer paths to reach them. Therefore, the objective to achieve lower power consumption leads to over-utilization of the network resources. On the other hand, the MNR objective achieves better network utilization, as it aims at minimizing the total path length. For high load values, all available DCs need to be powered up for both objectives. However, in this case, applying the MNR, achieves lower network resource utilization.

In Figure 4.4 we demonstrate the same set of results for a WDM network without wavelength conversion capability. The results acquired applying the two different objectives are similar as in the VWP case, demonstrating a similar trend, where the MJP achieves lower power consumption across most load values and the trend is changing only for the highest load value of 600 lightpaths. The difference of the absolute values is attributed to the absence of wavelength converters, which significantly reduces the network power consumption.

b) Aiming at a more detailed evaluation of the objective impact on the power consumption of the converged infrastructure, we provide a set of results that represent the total power consumption of the infrastructure across the VIs established over the physical infrastructure and for three different lightpath load volumes. Figure 4.7 and Figure 4.8 present these results for 30-60 and 120-150 lightpaths respectively for the virtual wavelength path case.

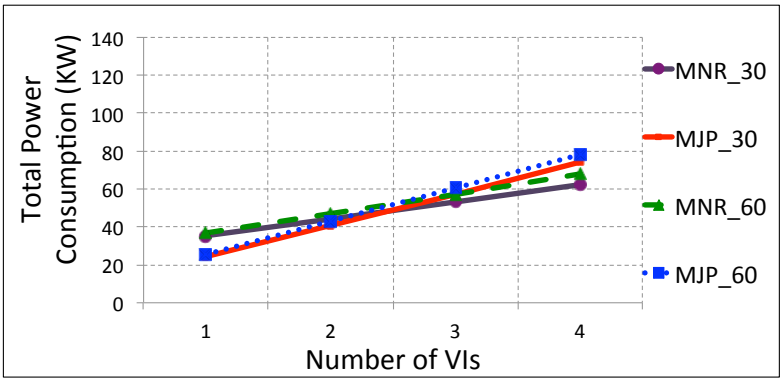

Figure 4.7 – Total Power Consumption vs Number of VIs - VWP 30-60 Lightpaths

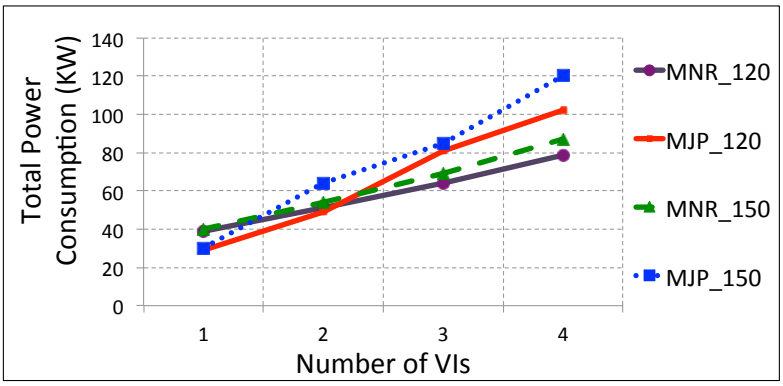

Figure 4.8 – Total Power Consumption vs Number of VIs - VWP 120-150 Lightpaths

The MJP objective achieves lower power consumption only for one or two planned VIs, whereas for higher number of VIs, MNR outperforms MJP for both low (30-60) and high (120-150) lightpath requests per VI. The observed change of trend is attributed to the fact that planning of multiple VIs is not based on a global optimization performed having in advance knowledge of the VI requests. Instead a practical planning approach is adopted, according to which, the planning procedure takes place for each VI request sequentially, considering that all previous VIs remain established and utilize the already assigned resources. Finally it should be noted that the addition of wavelength converters in the VWP case, assumed to employ OEO transponder technology in this study, significantly increases the overall power consumption of the optical network. This has an impact on the relative proportion of power consumption of the network and the data center resources and the associated trade-offs.

These observations clearly indicate that in order to maintain the energy efficiency benefit that MJP can offer when planning VIs over a PI, that are dynamically requested in time, there is a need to reconsider the existing resource allocation per VI periodically or following specific triggering events. This can then be followed by suitable reallocation of resources per VI through a VI re-planning phase. If there is a requirement for this reallocation of resources to not disrupt for services that are already supported, relevant constraints in the VI re-planning process can be applied.

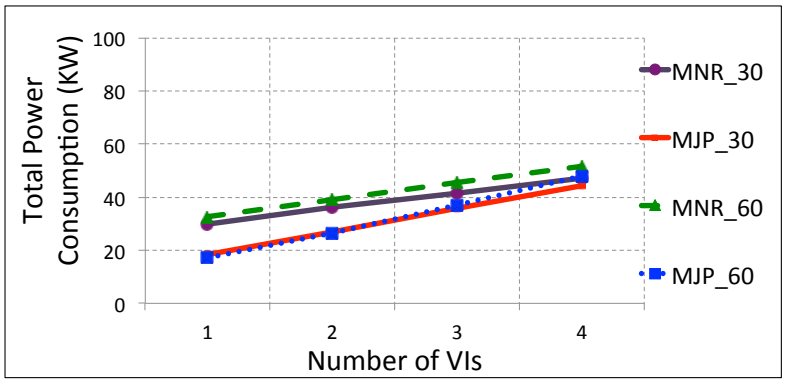

Figure 4.9 – Total Power Consumption vs Number of VIs - WP 30-60 Lightpaths

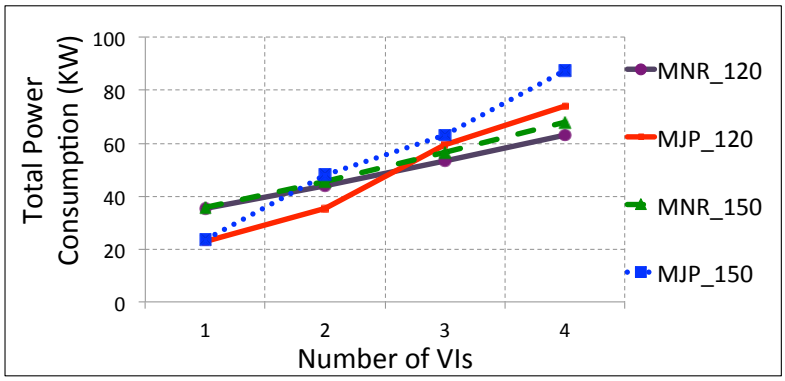

Figure 4.10 – Total Power Consumption vs Number of VIs - WP 120-150 Lightpaths

Figure 4.9 and Figure 4.10 provide the same set of results for the WP case, where wavelength converters are not present at the network nodes and the corresponding planning problem takes into account the fact that the same wavelength has to be assigned across each lightpath established on the physical topology, known as the wavelength continuity constraint. The graphs illustrating the power consumption for the different load values across the VIs are in accordance to the results discussed above and indicate similar performance for the two objectives and the WP case.

4.4.2 Blocking Performance

Figure 4.11 depicts the network resource utilization in terms of the number of link-wavelengths utilized by the VIs for the two objectives. The results are presented across

all lightpath volumes and after all four virtual infrastructures have been established. For the WP network, MNR achieves link-wavelength utilization of the order of 3%-13% compared to 8%-30% for the VWP network. MJP achieves a similar benefit, since no difference in the network resource assignment is observed between the two formulations for the specific problem.

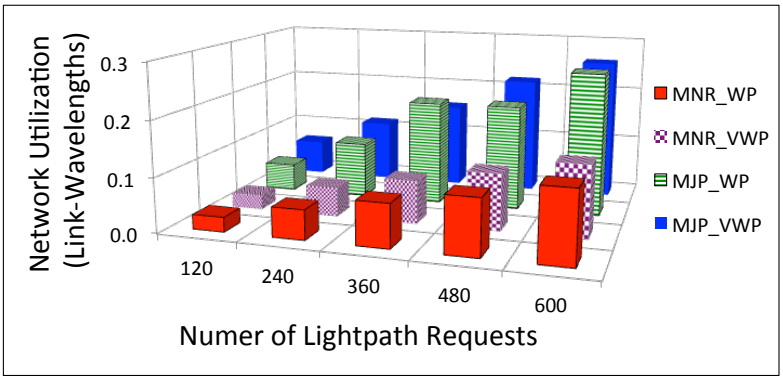

Figure 4.11 – Network Capacity Utilization

As already mentioned and observed through the investigation of power consumption contributions from the network and data center resources, the two objectives result in different virtual network infrastructures. The average number of wavelengths per fiber link is the same, but the node degree and total number of links used differs. MNR planned VIs exhibit a uniform node degree of 2 across all load volumes, whereas VIs created by the MJP objective exhibit node degrees of 3 due to the selection of longer paths.

Based on the VIs generated as described in detail above, we provide results that evaluate the performance of an online routing algorithm over the established planned VIs. We compare two virtual infrastructures differing in the node degree under both wavelength-converted and wavelength-continuity-constrained optical networks.

Figure 4.12 depicts the blocking performance of the two VIs in a wavelength-converted network (VWP) and quantifies the performance gain of the MJP category of VIs. The VI with node degree 3 (MJP) reaches a blocking probability value of 5%.The efficiency of

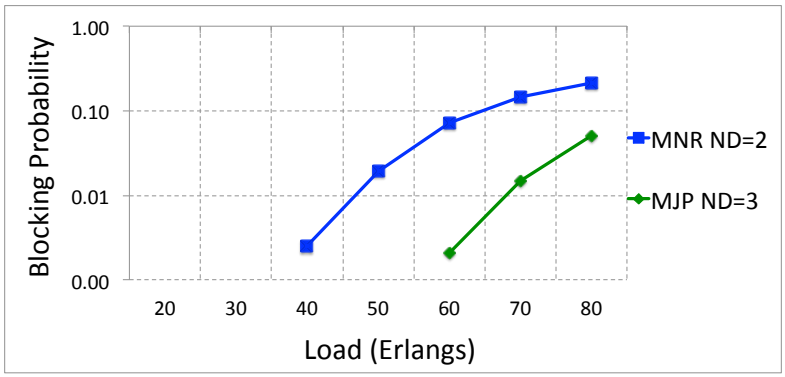

Figure 4.12 – Blocking Probablility vs Lightpath Requests - VWP

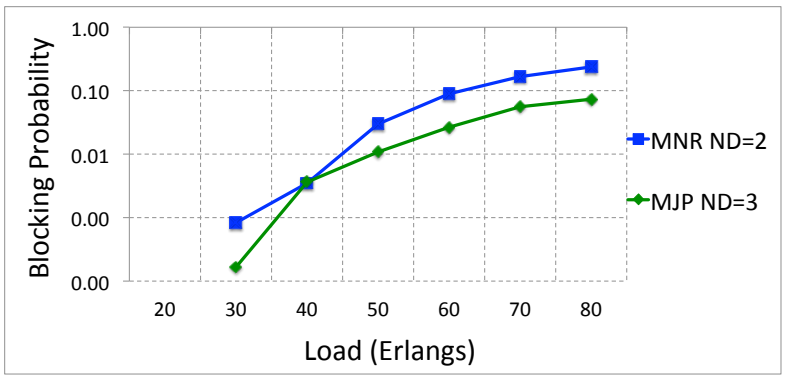

Figure 4.13 – Blocking Probablility vs Lightpath Requests - WP

the MNR objective in the utilization of network resources results in much higher blocking values that reach the level of 21% for the highest loading value. In the case where wavelength assignment is applied at each network node, both VI types experience higher demand blocking due to the increased probability of a path to be discarded because of the wavelength continuity constraint. Figure 4.13 illustrates blocking values of up to 7% for the MJP VI and 23% for the MNR VI. A much higher difference in the blocking probability between VWP and WP networks is observed for lower loading values and for both objectives.

4.5 Conclusions

In this chapter, we presented a detailed study of planning virtual infrastructures over a physical infrastructure comprising integrated optical network and data center resources. The study assumed a practical VI demand model that did not support any in advance global knowledge of the VI requests and through detailed ILP modeling compared two different objective functions the MJP and MNR as well as two different optical network architectures one supporting the VWP and one supporting the WP. The various scenarios under study were compared with regards to power consumption, network utilization and blocking performance of the planned VIs. Our results illustrated that although power consumption is an important aspect and an objective function (MJP) that optimizes the energy efficiency of the infrastructure, it may introduce inefficiencies in the utilization of network resources when the number of requests exceeds a certain level. This may in turn compromise the benefit with regards to energy efficiency, compared to what is achieved, when applying an objective that minimizes the network resource utilization for this high demand levels. To overcome this inefficiency, periodic re-planning of the requests can be applied. Finally, a set of dynamic traffic provisioning results were provided through simulations illustrating that the efficient resource utilization of the second objective (MNR) introduces a penalty on the produced VIs, especially in terms of connectivity, that leads to poor request blocking performance.

Chapter 5

Deployment of Correlated Cloud-Based Services

"Clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services). These resources can be dynamically re-configured to adjust to a variable load (scale), allowing also for an optimum resource utilization. This pool of resources is typically exploited by a pay-per-use model in which guarantees are offered by the Infrastructure Provider by means of customized SLAs " [4]. Among several early definitions present in the literature [74], [75], the one presented in [4] provides a simple but yet self-contained definition of Cloud computing as it has emerged during the recent years. The basis of this definition lies in some of the main functionalities/characteristics to support the Cloud paradigm, such as virtualization, dynamic reconfiguration, scalability, optimal resource utilization and pay-per-use model. Apart from these, the Cloud paradigm shares but is not limited to several other characteristics of Grid computing such as resource sharing, security, self-management and QoS guarantees, as summarized in [4].

The concept of a Virtual Machine (VM) [76] has been introduced to capture the use of virtualization technology to split the resources (computing, memory, storage, ...) of a single machine to several entities that run a separate operating system instance and enable isolation, consolidation and workload migration [77]. VM migration [76], [78] is

the implementation of workload migration that facilitates load balancing and maintenance purposes [77]. An important differentiation categorizes VM migration approaches into a) pure stop-and-copy or cold migration and b) live or hot migration [77], [79], [80]. In the former, VM is halted at the source, required action regarding memory copying at the destination is taken and then the VM is started at the destination. In the latter on the other hand, the Operating System (OS) and its applications are not interrupted and VM migration involves repeated data transfers that impose several drawbacks [77]. The main disadvantage of live VM migration lies on the service disruption in terms of downtime against the important benefits in terms of load balancing and fault tolerance.

The work presented in this chapter is motivated by [6] where the authors demonstrate that DC to client (D2C) cloud services present correlation patterns in terms of the need for interaction (data exchange) between pairs of services. These correlation patterns are not usually taken into account in the way that services are deployed across the multiple servers in different DCs in the form of VMs. Such services include Web 2.0, news, e-mail, video and gaming and the degree of need for interaction in [6] is illustrated in terms of pair-wise temporal port-based correlation based on the Pearson product-moment correlation coefficient. Based on these observations, we formulate and solve an optimal VM deployment problem for cross-correlated services among clients and DCs in a Cloud-based infrastructure. The relevant problems do not take into account the Cloud-related implementation details, but are rather presented as design/planning problems in the form of Integer Linear Programming, aiming to act as the basis for identifying trends and proposing efficient deployment solutions. In this context, we model correlation as a random process that follows a known probability distribution and each service pair has a known correlation coefficient. For this reason we employ stochastic optimization tools and more specifically the Sample Average Approximation (SAA) method to solve the relevant problems. To illustrate the benefits of the stochastic optimization approach and also the impact of correlation on the excess capacity and cost incurred, we also solve appropriate problems that treat correlation as a deterministic input and problems that do not consider correlation respectively.

Service deployment is modeled in terms of VMs running on Data Centers (DCs) and migration of these VMs from one DC to another is considered where applicable. The main goal is to demonstrate the impact that service correlation has on the network and DC resource utilization and thus the total cost of the cloud infrastructure. We refer to the service requested by the client as the first-stage service and to the DC where the respective VM is deployed as the first-stage DC. Accordingly, the service that the first-stage service is correlated with is reffered to as the second-stage service and the relevant DC as the second-stage DC. The impact of correlation relies on the need for a first-stage service to communicate with the correlated second-stage service in order to exchange data before the user is served. We can easily understand that this can have an important impact on resource utilization and service provisioning depending on how the second stage communication is implemented on the one hand and on how the services are deployed on the other hand. Regarding the stochastic version of the problem, we consider multiple service groups with different correlation coefficient characterizing the pair of services in each group. Correlation per service pair is modeled as a pair of real numbers that follow a known probability distribution. Second-stage communication is implemented either in terms of VM migration or in terms of a DC to DC (D2D) path and is based on a comparison rule that involves a pre-defined threshold.

In order to illustrate all these cases, we model two kinds of problems: a) a Deterministic Problem (DP), where for each service request a first-stage path and a VM on one of the available DCs are assigned, whereas second-stage communication is established where applicable between the service requested and the corresponding correlated service on a distinct DC. Various flavors of this problem are examined: a-i) the extreme case where correlation is not taken into account at all and thus only first-stage routing and VM deployment decisions are made, the case where correlation is considered and second-stage decision involves a-ii) either a path assignment from the first-stage DC to another DC where the correlated service is deployed, or a-iii) the migration of this VM for the second-stage service from the second-stage to the first-stage DC. The two last cases are modeled as worst cases and thus all services are considered to require the second-stage

communication due to the correlation. b) a Stochastic Problem (SP), where the uncertainty lies in the correlation of service pairs and the Sample Average Approximation (SAA) method is employed to provide us with solution estimates. In all cases, costs are assigned for the establishment of ligthtpaths, the VM deployment and VM migration and the total cost minimization is the objective of all problems. The results illustrate a) the impact of second-stage communication approach on the deterministic problems in terms of the increased cost required due to the additional network and DC capacity deployed and b) the important cost savings offered by the stochastic correlation modeling and appropriate VM deployment. Finally, we present a significant decision against which the infrastructure provider is faced in the existence of an upper bound on cost regarding the choice of the optimal threshold that defines whether service correlation will lead to second-stage communication.

5.1 Related Work

In [4], the authors present the actors, such as service providers and infrastructure providers, and the scenarios, such as Infrastructure as a Service (Iaas), Platform as a Service (PaaS) and Software as a Service (SaaS) that form the basic Cloud system. Based on these and after summarizing the most important definitions of cloud computing, they propose a Cloud definition whose main focusing is on characteristics such as "scalability, pay-per use model and virtualization". Finally, they provide a comparison of the main features of Grid and Cloud such as resource sharing, virtualization, scalability etc., aming at a clear distinction between the two. In [77] the authors position VM migration in the context of the Cloud paradigm and identify it as a key basic enabling technology to ensure important features such as scalability and load balancing. After differentiating it with "pure stop-and-copy" or "hot" migration, they present basic characteristics of modern Internet applications that require the advanced performance benefits offered by live migration. Based on this motivation, they implement benchmarking experiments over realistic workloads aiming at providing a more clear view of live VM

migration. The results obtained over a Web 2.0 application example demonstrate significant service disruption in terms of downtime, making the development of more efficient solutions required for the benefits of live VM migration to be fully utilized. In [81] the authors present a distributed approach for solving the dynamic service placement problem with the objective to minimize "response time violations and resource rental cost at run time". In [82], the authors deal with the problem of "service co-deployment"in cloud computing by formulating an integer program that "minimizes the latency of user requests". They first provide a model for single service deployment where users are assigned to be served based on the distance criterion and then extend this model to a multi-service environment, where correlation between services is taken into account. In this case, the solution of the problem provides the co-deployment of related services for each request at once and aims to minimize the total latency experienced by the user. To mitigate the huge problem size, the authors also provide an iterative heuristic algorithm for the proposed deployment scheme. Based on real-world latency data collected for the purpose of this work, they provide experimental results on latency and demonstrate significant savings compared to the independent deployment scheme where the relation of services is not taken into account. [83] presents how computational, storage and network resources can be efficiently utilized through a network-aware cloud. Through a prototype demonstration, the authors demonstrate how efficient service placement and migration among distributed data centers can mitigate the user experience degradation due to any possible cloud infrastructure issues. By comparing the distributed algorithm with a greedy centralized algorithm, they indicate that near-optimal solutions are produced by the proposed approach under static traffic. Finally, they present results under dynamic traffic, also compared to the centralized approach at specific time instances and demonstrate adaptivity to the traffic pattern and improved performance compared to the centralized solution in some cases.

5.2 Methodology Background

According to [84], "the aim of stochastic programming is to find an optimal decision in problems involving uncertain data", as opposed to deterministic programming, where uncertainty is not taken into account. In this chapter, the related problem is formulated as a two-stage stochastic routing problem, as presented in detail in [25]. In two-stage stochastic routing problems, the first stage involves a route selection and the second stage some recourse, that being a re-routing decision, a penalty or any decision affected by the first stage. The optimization objective is expressed as the total cost incurred by both the first stage decisions and the expected cost of the recourse. The general class of two stage stochastic routing problems (also referred to as the original problem) can be formulated as in Eq. (5.1)

$$z* = \min\ c^T x + E_P\left[Q(x, \xi(\omega))\right] \tag{5.1}$$

where

$$Q(x, \xi(\omega)) = \min\ _{y \geq 0}\{q(\omega)^T y | Dy \geq T(\omega)x\} \tag{5.2}$$

and x is the first stage routing decision, y is the second stage recourse decision and $\omega \in \Omega$ is the scenario where uncertainty lies in and is not known during the first stage decision but its probability distribution is considered known during the second stage decision.

Several solution approaches exist in the literature regarding stochastic routing problems. These are based on heuristics [85], dynamic programming [86], or exact methods for cases with a small number of scenarios using the L-shaped method [87]. For problems with very large number of scenarios, exact evaluation of the second stage expected value function in Eq. (5.1) is impossible, thus several sampling methods have been proposed in the literature [25]. Among them, the Sample Average Approximation (SAA) method is an exterior sampling method in which the expected value function $E_P[Q(x, \xi(\omega))]$ is approximated by the sample average function $\sum_{n=1}^{N} \frac{1}{N} Q(x, \xi(\omega))$.

The SAA problem is formulated as

$$z_n = \min\, c^T x + \frac{1}{N}\sum_{n=1}^{N} Q(x, \xi(\omega^n)) \tag{5.3}$$

The optimal objective value and optimal solution of the original problem 5.1 are estimated by the optimal value z_n and an optimal solution $\hat{x}$ obtained by solving the SAA problem.

By generating M independent samples each of size N and solving the associated SAA problems 5.3, we obtain values $z_N^1, ..., z_N^M$ and candidate solutions $\hat{x}^1, ..., \hat{x}^M$.

Let

$$\bar{z}_N = \frac{1}{M}\sum_{m=1}^{M} z_N^M \tag{5.4}$$

denote the average of the optimal objective values of the M SAA problems. Then $E\left[\bar{z}_N\right] \leq z^*$, thus $\bar{z}_N$ provides a statistical estimate for a lower bound on the optimal value of the true problem.

For any feasible point $\hat{x} \in X$, the objective value $c^T\hat{x} + E\left[Q(\hat{x}, \xi(\omega))\right]$ is an upper bound for z^*. The upper bound can be estimated by

$$\hat{z}_{N'}(\hat{x}) = c^T\hat{x} + \frac{1}{N'}\sum_{n=1}^{N'} Q\left(\hat{x}, \xi(\omega^n)\right) \tag{5.5}$$

where $\{\omega^1, \omega^2, ..., \omega^{N'}\}$ is a sample of size N'. Note that $N' > N$ and that the sample of size N' is independent from the sample used to generate $\hat{x}$. Then $\hat{z}_{N'}(\hat{x})$ is an unbiased estimator of $c^T\hat{(}x) + E\left[Q(x, \xi(\omega))\right]$ and hence for any feasible solution $\hat{x}$:

$$E\left[\hat{z}_{N'}(\hat{x})\right] \geq z^* \tag{5.6}$$

From the M optimal solutions $\hat{x}^1, \hat{x}^2, ..., \hat{x}^M$ produced, we choose the one that has the smallest objective value:

$$\hat{x}^* \in argmin\{\hat{z}_{N'}(\hat{x}) : \hat{x} \in \{\hat{x}^1, \hat{x}^2, ..., \hat{x}^M\}\} \tag{5.7}$$

The optimality gap estimate is calculated as

$$\hat{z}_{N'}(\hat{x}^*) - \bar{z}_N \tag{5.8}$$

where $\hat{z}_{N'}(\hat{x}^*)$ is re-computed after performing the minimization in 5.7 with an independent sample to obtain an unbiased estimate.

5.3 Problem Formulation

In terms of the stochastic problem formulation, service correlation is modeled as follows: Each traffic demand is assumed to require one wavelength in terms of network capacity and one VM in terms of DC resources. All services are assumed to require the same amount of computing, storage and memory resources and thus VMs requirements are the same across services. Each demand requires one service and this is called the first-stage service. The offered services are classified in two groups in the case of stochastic problems: a) a highly correlated group and b) a low correlated group. Each group consists of a service pair and services in each group are indexed by $a \in A$, whereas groups are indexed by $g \in G$ and demands by $d \in D$. We use a_d to refer to the service that the client asks in the first-stage and g_d to the group that this demand belongs to.

The correlation of a service pair is defined as the Pearson product-moment correlation coefficient ρ. According to the service group that a pair belongs to, we use the corresponding correlation coefficient to generate a two-dimensional random vector $c = (c_1, c_2)$ whose values correspond to the pair of services of the group. For this we assume a bivariate normal distribution, so that the probability density function of the vector c is given by Eq. (5.9).

$$f(c_1, c_2) = \frac{1}{2\pi\sigma_{c_1}\sigma_{c_2}\sqrt{1-\rho^2}} exp\left(-\frac{1}{2(1-\rho^2)} \left[\frac{(c_1-\mu_{c_1})^2}{\sigma_{c_1}^2} + \frac{(c_2-\mu_{c_2})^2}{\sigma_{c_2}^2} - \frac{2\rho\,(c_1-\mu_{c_1})\,(c_2-\mu_{c_2})}{\sigma_{c_1}\sigma_{c_2}}\right]\right) \tag{5.9}$$

where

$\sigma_{c_1} > 0$, $\sigma_{c_2} > 0$, $\mu = \begin{pmatrix} \mu_{c_1} \\ \mu_{c_2} \end{pmatrix}$ and $\Sigma = \begin{pmatrix} \sigma_{c_1}^2 & \rho\sigma_{c_1}\sigma_{c_2} \\ \rho\sigma_{c_1}\sigma_{c_2} & \sigma_{c_1}^2 \end{pmatrix}$ is the variance-covariance matrix.

Upon generation of the random values for each scenario according to the SAA method, we have a pair of values (c_1, c_2) for each demand that is generated based on the correlation coefficient of the group that the services belong to. The use of the two-dimensional matrix is two-fold: 1) among the two services, the one with the largest value is selected to be the first-stage service that is the one initially required by the client, denoted by a_d. 2) it enables us to establish whether second-stage communication is required. This is supported through the calculation of the absolute difference of the two values, that is $|c_1 - c_2|$. If $|c_1 - c_2| < c_{th}$, then service a_d needs to communicate with the other service of the group from a different DC, whereas if $|c_1 - c_2| \geq c_{th}$, communication is not required. The initial assignment of VMs on DCs is only in terms of capacity for each group of services and no deployment of services is considered. Thus for every second-stage communication required by correlation of the service pair, the solution of the problem gives us either the distinct DC that the secondary service is deployed and the corresponding D2D path, or the migration of a VM for this service from a distinct DC to the first-stage DC, always obeying the capacity bounds of the DCs for each group of services. This way, we do not consider the case where enough VM capacity exists on the first-stage DC and no migration or D2D path is needed for communication with another DC, although the second-stage communication is required. Each demand is characterized by the triplet $\{o, c_1, c_2\}$, where o is the origin node selected from a uniform distribution among the available client nodes of the network.

We formulate, solve and compare two different classes of problems: a) the class of Deterministic Problems (DP) that consists of a problem that does not consider any correlation between services (DP_{noCor}) and two worst-case problems where all services are considered to be fully correlated with another service of the same group. These two problems differ in the second-stage communication between the correlated services, since in one case migration of a VM for the correlated service takes place from a different

DC to the first-stage DC (DP_{mig}) (live migration) and in the other case VM migartion is not allowed but rather a D2D path has to be established (DP_{d2d}). b) the class of Stochastic Problems that consists of three problems. These differ in the percentage of high ($\rho = 0.9$) and low ($\rho = 0.1$) correlated traffic that each one consists of. The first, SP_{low} consists of 90% low correlated traffic and 10% highly correlated traffic, the second SP_{high} consists of 90% highly and 10% low correlated traffic and the third, SP_50_50 consists of the same amount (50%) of highly and low correlated traffic. A mean value of 5 and a variance of 0.2 are used. The absolute difference of the values generated by the bivariate normal distribution gets higher as the correlation coefficient gets lower and thus the probability for second-stage communication to be required gets smaller as the threshold value gets smaller. The threshold values against which we compare the outcome of the random values are $(0.1, 0.2, 0.3, 0.4, 0.5)$. The problems are summarized in Table 5.1. [1].

Table 5.1 – Problems

Problem	Uncertainty	Correlation	2^{nd} stage communication
DP_{noCor}	no	no	no
DP_{mig}	no	full	live migration
DP_{d2d}	no	full	D2D path
SP_{low}	yes	random	migration or D2D path
SP_{5050}	yes	random	migration or D2D path

The converged network and DC infrastructure is modelled as a graph $G = (N, E)$ where N is the set of nodes and E is the set of edges (links) interconnecting these nodes. All network nodes $n \in N$ are modelled as an optical cross connect (OXC) Figure 3.1. S is a subset of N and depicts the set of nodes where DCs are attached. The set $\{N - S\}$ represents the client nodes where demands are initiated. Each link interconnecting the nodes is a fiber link with a maximum capacity of 100 wavelengths and one such directional link is assumed between any two interconnected nodes. DC s

[1] SP_* refers to "low","5050" and "high" stochastic problems

capacity in terms of VMs is also upper bounded by V_s^{max} which is 60 for each DC in our case. All problems are solved for total traffic volumes of 20, 40 and 60 lightpath/VM requests and result in an optimal VM and network capacity placement on the available DCs and WDM network links respectively. The developed algorithms aim to tackle the problem of optimal VM deployment for a set of requests and thus have a planning/design nature. For this reason, they are formulated and solved in terms of binary Integer Linear Programs (ILP). Table 5.2 includes the inputs and parameters used throughout the different problem formulations.

Table 5.2 – Problem Parameters

Parameter	Description
H_D	total demands volume in terms of VMs (lightpaths)
h_d	volume of demand d
$\|\|S\|\|$	total number of DCs
V_s^{max}	maximum capacity of DC s in terms of VMs (lightpaths)
δ_{edsp}	=1, if candidate path p for demand d towards DC s contains link e
$\gamma_{edss'q}$	=1, if candidate path q for demand d for D2D communication between 1^{st} stage DC s and 2^{nd} stage DC s' contains link e
c^{th}	threshold value for the absolute difference cross-correlation values between services in the same group
b_a^d	=1, if service a of group g_d needs to communicate with service a_d. This is the result of the randomly generated pair of values from the bivariate normal distribution with the correlation coefficient of the group g as an input.

The set of variables used across the problem formulations are presented in Table 5.3. Regarding the flow variables for 1^{st} stage (x_{dsp}) and second-stage ($z_{dss'q}$) communication paths, we employ Yen's k-shortest path routing algorithm with $k = 2$ to populate the list of candidate paths for each demand from the client node towards any of the available DC destinations and from the 1^{st} stage DC to the 2^{nd} stage DC.

Table 5.3 – Problem Variables

Variable	Domain	Description
x_{dsp}	$\{0,1\}$	=1, if candidate path p is used to realized demand d towards DC s
$z_{dss'q}$	$\{0,1\}$	=1, if candidate path q is used to realize demand d's need for D2D communication between DC s and DC s'
$m^{da}_{ss'}$	$\{0,1\}$	=1, if VM serving service a (of the same group g_d) for demand d migrates from DC s to DC s'
$z^{da}_{ss'q}$	$\{0,1\}$	=1, if candidate D2D path q for service a (in the same group d_g) of demand d from DC s to DC s' is established
$y^{da}_{ss'q}$	$\{0,1\}$	Used for linearization due to the multiplication of $m^{da}_{s's}$ and $z_{dss'q}$
w_e	Z_+	number of wavelengths established on link e
v^{ga}_s	Z_+	number of VMs for service a of group g running on DC s

In the following subsections we present the formulations of the stochastic (Section 5.3.1) and the deterministic problems (Section 5.3.2) based on the parameters and variables presented above.

5.3.1 Stochastic Problem

Throughout the problem formulations, we use the common objective function presented in Eq. (5.10) that represents the total cost of VM deployment across the DCs and lightpaths establishment over the fiber links interconnecting the DCs. The relevant costs of VM hosting on a DC, VM migration from a DC to another and lightpath establishment are presented in Table 5.4.

Table 5.4 – Cost Values

Symbol	Description	Value
C^{VM}_{dep}	Cost of deploying a VM on a DC	\$0.1 per VM (per period)
C^{VM}_{mig}	Cost of VM migration	\$0.01 per VM (per period)
C_{path}	Cost of path establishment	\$0.2 per VM (per GB)

The objective function for the 2-stage stochastic problem consists of two main parts: 1) the cost of first-stage VM deployment illustrated by the first-stage variable x_{dsp} and including the cost of establishing the path C_{path} from the client to the first-stage DC (D2C path) and the cost of deploying the VM for the first-stage service a_d at the first-stage DC, C_{dep}^{VM}. 2) The cost of the recourse in terms of the stochastic optimization. This second part is illustrated through the variables of VM migration $m_{s's}^{da}$ and D2D path $z_{ss'q}^{da}$ and the respective costs (C_{mig}^{VM}, C_{dep}^{VM}) and (C_{path}, C_{dep}^{VM}). We should note that the cost of VM hosting is also considered in the cases of VM migration and/or D2D path establishment. For the stochastic problems, second-stage communication initiated due to the correlation of service pairs is implemented either through the migration of a VM from a DC (referred to also as second-stage) to the first-stage DC where service a_d is already deployed, or through the establishment of a D2D path again between the first-stage DC and a distinct DC with available capacity. The mutual exclusion of these two choices is illustrated in the objective function through the multiplication of each of the respective variables with the complementary of the second. We thus introduce a third binary variable $z_{ss'q}^{da}$ to replace this multiplication and mitigate the non-linearity. The following formulation is the same for all three stochastic problems that differ only in the mixture of traffic regarding the percentage of highly and low correlated traffic.

$$
\begin{aligned}
min\ C_{tot} =& c^T x + E_P\left[Q\left((z,m),\xi(\omega)\right)\right] \\
=& \sum_d \sum_s \sum_p \left(C_{path} + C_{dep}^{VM}\right) x_{dsp} + \\
& \frac{1}{N}\left\{ \sum_d \sum_{\substack{s,s' \\ s\neq s'}} \sum_{\substack{a, \\ a\neq a_d}} b_a^d \left[\left(C_{mig}^{VM} + C_{dep}^{VM}\right) m_{s's}^{da} \sum_q \left(1 - z_{dss'q}^{da}\right) + \right.\right. \\
& \left.\left. \left(1 - m_{s's}^{da}\right) \sum_q z_{dss'q}^{da} \left(C_{depl}^{VM} + C_{path}\right) \right] \right\}
\end{aligned} \tag{5.10}
$$

which after the linearization becomes:

$$
\begin{aligned}
min\ C_{tot} =& c^T x + E_P\left[Q\left(\{z, m\}, \xi(\omega)\right)\right] \\
=& \sum_d \sum_s \sum_p \left(C_{path} + C_{dep}^{VM}\right) x_{dsp} + \\
& \frac{1}{N} \Bigg\{ \sum_d \sum_{\substack{s,s' \\ s \neq s'}} \sum_{\substack{a, \\ a \neq a_d}} b_a^d \Bigg[\left(C_{mig}^{VM} + C_{dep}^{VM}\right) m_{s's}^{da} \sum_q \left(1 - y_{dss'q}^{da}\right) + \\
& \sum_q \left(C_{dep}^{VM} + C_{path}\right) \left(z_{dss'q}^{da} - y_{dss'q}^{da}\right) \Bigg] \Bigg\}
\end{aligned} \tag{5.11}
$$

The cost of deploying the VMs to the available DCs and establishing the lightpaths over the fiber links in order to accommodate the client requests for correlated services is minimized subject to the set of constraints presented below.

All demands served : Ensure that all demands are accommodated, both for first-stage and second-stage communication through the variables x_{dsp} and $m_{s's}^{da}$ (migration) or $z_{ss'q}^{da}$ (D2D path).

$$
\sum_s \sum_p x_{dsp} + \sum_{\substack{s,s' \\ s \neq s'}} \sum_{\substack{a \\ a \neq a_d}} b_a^d \left(m_{s's}^{da} + sum_q z_{ss'q}^{da}\right) \geq h_d, d = 1, ..., D \tag{5.12}
$$

Link Capacity : Ensure that the sum of first-stage and second-stage flows utlizing link e is less that the total link capacity w_e

$$
\sum_d \sum_s \sum_p \delta_{edsp} x_{dsp} + \sum_d \sum_{\substack{s,s' \\ s \neq s'}} \sum_{\substack{a \\ a \neq a_d}} b_a^d \gamma_{edss'q}^{da} z_{ss'q}^{da} \leq w_e, e = 1, ..., E \tag{5.13}
$$

DC s capacity for each service : The total number of VMs for service a of group g on DC s has to be enough to accommodate the VMs for first-stage services plus the VMs of second-stage services, both those that are migrated from any DC s' to DC

s and those that are offered from DC s through a D2D path.

$$\sum_{\substack{d \\ a_d=a \\ g_d=g}} \sum_{p} x_{dsp} + \sum_{\substack{d \\ a_d \neq a \\ g_d=g}} \sum_{\substack{s' \\ s' \neq s}} b_a^d \left(m_{s's}^{da} + \sum_{q} z_{s'sq}^{da} \right) \leq u_s^{ga}, s = 1, ..., S \tag{5.14}$$

Upper bound on VMs running on DC s : The total number of VMs deployed at DC s for all services should be less than the total DC capacity in terms of DCs.

$$\sum_{g} \sum_{a} u_s^{ga} \leq V_s^{max}, s = 1, ..., S, g = 1, ..., G, a = 1, ..., A \tag{5.15}$$

VM migration origin/destination constraints (part a) : VM migration for second-stage communication cannot be originated at the DC s where first-stage service a_d is served.

$$\sum_{p} x_{dsp} + \sum_{\substack{a \\ a \neq a_d}} \sum_{\substack{s' \\ s' \neq s}} b_a^d m_{ss'}^{da} \leq 1, d = 1, ..., D, s = 1, ..., S \tag{5.16}$$

VM migration origin/destination constraints (part b) : The VM migration destination for second-stage communication cannot take place between two DCs that are both different than the first-stage DC s.

$$\sum_{p} x_{dsp} + \sum_{\substack{a \\ a \neq a_d}} \sum_{\substack{s' \\ s' \neq s}} \sum_{\substack{s'' \\ s'' \neq s' \\ s'' \neq s}} b_a^d m_{s's''}^{da} \leq 1, d = 1, ..., D, s = 1, ..., S \tag{5.17}$$

D2D paths constraints (part a) : The D2D path for second-stage communication originates from the first-stage DC, not from the second-stage (used for uniformity on how paths are assigned).

$$\sum_{p} x_{dsp} + \sum_{\substack{a \\ a \neq a_d}} \sum_{\substack{s' \\ s' \neq s}} \sum_{q} b_a^d z_{s'sq}^{da} \leq 1 \tag{5.18}$$

D2D paths constraints (part b) : The D2D path for second-stage communication has to

include the first-stage DC and a distinct DC.

$$\sum_{p} x_{dsp} + \sum_{\substack{a \\ a \neq a_d}} \sum_{\substack{s' \\ s' \neq s}} \sum_{\substack{s'' \\ s'' \neq s' \\ s'' \neq s}} \sum_{q} b_a^d z_{s's''q}^{da} \leq 1 \tag{5.19}$$

Linearization : Set of constraints used to linearize the objective function due to the multiplication of the binary variables $m_{s's}^{da}$ and $z_{dss'q}^{da}$ through the binary variable $y_{ss'q}^{da}$

$$\begin{aligned} y_{ss'q}^{da} &\leq m_{s's}^{da} \\ y_{ss'q}^{da} &\leq z_{dss'q}^{da} \\ y_{ss'q}^{da} &\geq m_{s's}^{da} + z_{dss'q}^{da} - 1 \end{aligned} \tag{5.20}$$

5.3.2 Deterministic Problems

The general formulation for the class of deterministic problems is presented in this subsection. The same sets of variables and parameters with the stochatic formulation are used here, apart from the parameter b_d^a that reflects the uncertainty of the need for second-stage communication for service a of demand d due to correlation with the other service of the group g_d. As mentioned before, we solve three deterministic problems, DP_{mig}, DP_{d2d} and DP_{noCor}. Specifically for the first two problems where correlation is taken into account and treated with VM migration and D2D path establishment respectively, we consider a deterministic and worst-case scenario for correlation between each service pair, meaning that second-stage communication is required for the services of every demand. The formulation of each problem falls under the generic ILP formulation presented here, which however includes some modifications required to address the details corresponding to each specific case. The differences of each problem are listed below:

DP_{mig} : Second-stage communication is realized through VM migration from a distinct DC with available capacity for the specific service to the first-stage DC that serves

the client request. This approach can be classified as live VM migration with more stringent requirements in terms of delay, synchronization and implementation cost. However, the cost of deploying the VM and VM migration used in the set of problems presented here is the same.

DP_{d2d} : Second-stage communication is realized through the establishment of a D2D path from a distinct DC with available capacity for the specific service correlated with the client-requested service a_d that is deployed at the first-stage DC. No VM migration is considered in this case. Due to the high cost of D2D path establishment compared to the cost of VM migration, we expect this problem to provide a worst-case upper bound for all problem solutions and especially for the stochastic problems.

DP_{noCor} : Correlation of services is not considered at all, thus only first-stage VM deployment and bandwidth assignment form the problem solution. This problem is expected to provide a lower bound for the rest of the problems since no VM migration, deployment or second-stage services and D2D paths are considered.

$$
\begin{aligned}
min\ C_{tot} = & \sum_{d}\sum_{s}\sum_{p}\left(C_{path}+C_{dep}^{VM}\right)x_{dsp}+ \\
& \sum_{d}\sum_{\substack{s,s'\\ s\neq s'}}\sum_{\substack{a,\\ a\neq a_d}}\left[m_{s's}^{da}\left(C_{mig}^{VM}+C_{dep}^{VM}\right)\sum_{q}\left(1-y_{dss'q}^{da}\right)+\right. \\
& \left.\sum_{q}\left(C_{dep}^{VM}+C_{path}\right)\left(z_{dss'q}^{da}-y_{dss'q}^{da}\right)\right] \qquad (5.21)
\end{aligned}
$$

All demands served :

$$
\sum_{s}\sum_{p}x_{dsp}+\sum_{\substack{s,s'\\ s\neq s'}}\sum_{\substack{a\\ a\neq a_d}}\left(m_{s's}^{da}+\sum_{q}z_{ss'q}^{da}\right)\geq h_d, d=1,...,D \qquad (5.22)
$$

Link Capacity :

$$\sum_{d}\sum_{s}\sum_{p}\delta_{edsp}x_{dsp}+\sum_{d}\sum_{\substack{s,s'\\ s\neq s'}}\sum_{\substack{a\\ a\neq a_d}}\gamma^{da}_{edss'q}z^{da}_{ss'q}\leq w_e, e=1,...,E \tag{5.23}$$

DC s capacity for each service :

$$\sum_{\substack{d\\ a_d=a\\ g_d=g}}\sum_{p}x_{dsp}+\sum_{\substack{d\\ a_d\neq a\\ g_d=g}}\sum_{\substack{s'\\ s'\neq s}}\left(m^{da}_{s's}+\sum_{q}z^{da}_{s'sq}\right)\leq u^{ga}_{s}, s=1,...,S \tag{5.24}$$

Upper bound on VMs running on DC s

$$\sum_{g}\sum_{a}u^{ga}_{s}\leq V^{max}_{s}, s=1,...,S, g=1,...,G, a=1,...,A \tag{5.25}$$

VM migration origin/destination constraints (part a)

$$\sum_{p}x_{dsp}+\sum_{\substack{a\\ a\neq a_d}}\sum_{\substack{s'\\ s'\neq s}}m^{da}_{ss'}\leq 1, d=1,...,D, s=1,...,S \tag{5.26}$$

VM migration origin/destination constraints (part b)

$$\sum_{p}x_{dsp}+\sum_{\substack{a\\ a\neq a_d}}\sum_{\substack{s'\\ s'\neq s}}\sum_{\substack{s''\\ s''\neq s'\\ s''\neq s}}m^{da}_{s's''}\leq 1, d=1,...,D, s=1,...,S \tag{5.27}$$

D2D paths constraints (part a)

$$\sum_{p}x_{dsp}+\sum_{\substack{a\\ a\neq a_d}}\sum_{\substack{s'\\ s'\neq s}}\sum_{q}z^{da}_{s'sq}\leq 1 \tag{5.28}$$

D2D paths constraints (part b)

$$\sum_{p} x_{dsp} + \sum_{\substack{a \\ a \neq a_d}} \sum_{\substack{s' \\ s' \neq s}} \sum_{\substack{s'' \\ s'' \neq s' \\ s'' \neq s}} \sum_{q} z^{da}_{s's''q} \leq 1 \tag{5.29}$$

Linearization

$$\begin{aligned} y^{da}_{ss'q} &\leq m^{da}_{s's} \\ y^{da}_{ss'q} &\leq z^{da}_{dss'q} \\ y^{da}_{ss'q} &\geq m^{da}_{s's} + z^{da}_{dss'q} - 1 \end{aligned} \tag{5.30}$$

5.4 Results and Discussion

In this section we present the results obtained after solving the six distinct problems described in Table 5.1. All problems are solved for the converged optical network and DC infrastructure with 4 DCs deployed across the COST239 [39] reference network, as this is presented in Figure 4.2. The fiber capacity for each link is 100 wavelengths and each DC supports 60 VMs in total, uniformly spread across the two service groups and the pair of services for each group, where applicable. Since differentiation of traffic in terms of correlation groups (highly and low correlated group) is considered only in the stochastic problems, the total DC capacity is also uniformly spread across the two groups and among services of each group. This means that for each service in the pair of services of each group, 15 VMs are considered as the maximum capacity for each DC.

We present results for the three deterministic problems DP_{mig}, DP_{d2d} and DP_{noCor} as these are presented in the previous subsection and for three versions of the stochastic problem: a) SP_{low}, b) SP_{5050} and c) SP_{high}. Low and high refer to the percentage of low and correlated traffic respectively. Low correlated traffic is modelled as a group of two services with correlation coefficient equal to 0.3, whereas a value of 0.9 is used for the highly correlated traffic group. The percentage of the dominant traffic is 90% in each of

these two cases. All problems are solved for 20, 40 and 60 total volume of lightpath/VM requests. Specifically for the stochastic problems, the presented results are obtained for 5 samples of size 50 for the lower bound and 500 for the upper bound. The result regarding the optimality gap estimate that reflects how the solution estimation quality gets better across increasing number of samples is produced for 5, 10 and 20 samples, each of size 50 for the lower bound and 500 for the upper bound. Also, a lower number of demands (10) has been considered, since the increased number of considered scenarios also increases computational complexity.

Most of the results presented in this work are related to the correlation threshold c_{th} as this is defined in Section 5.3. Figure 5.1 - Figure 5.3 illustrate the total cost of deploying the required VMs across the 4 DCs and assigning the required capacity over the fiber links for 20, 40 and 60 lightpath requests originating from all client nodes. For all three cases, we observe the lower bound provided by the DP_{noCor} problem and the upper bound provided by the DP_{mig} problem, since all solutions of the stochastic problems lie between the two. We also observe the worst-case cost illustrated by the DP_{d2d} problem, where migration of VMs is not permitted and D2D paths have to be established. Across the deterministic problems, the results illustrate a 36% increase for the DP_{mig} problem and a 100% increase for the DP_{d2d} problem in the total cost compared to the DP_{noCor} problem. This way we identify a) the impact of service correlation on the cost of deploying the required VMs and assigning lightpaths under two different implementation approaches, VM migration and D2D paths and b) the cost savings offered, even in the case of approaching service cross-correlation in a deterministic way, by enabling migration of VMs instead of establishing inter-DC paths, and the analogy of these savings to the respective costs of VM migration and path establishment. The latter observation illustrates the importance of live migration and the need to implement efficient algorithms that provide minimum delay when pre-planned migration is not an option.

The three stochastic problems are differentiated by varying the percentage of highly and low correlated traffic and a cost difference that lies between 5% and 8% is observed

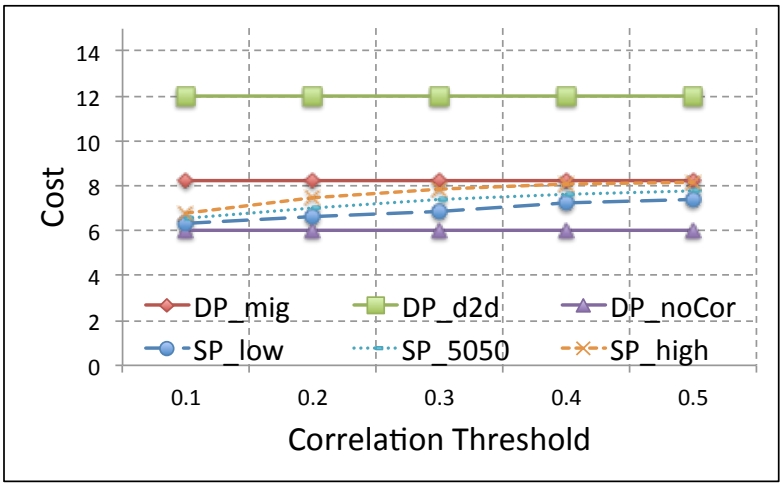

Figure 5.1 – Total Cost vs Correlation Threshold - 20 lightpath/VM requests

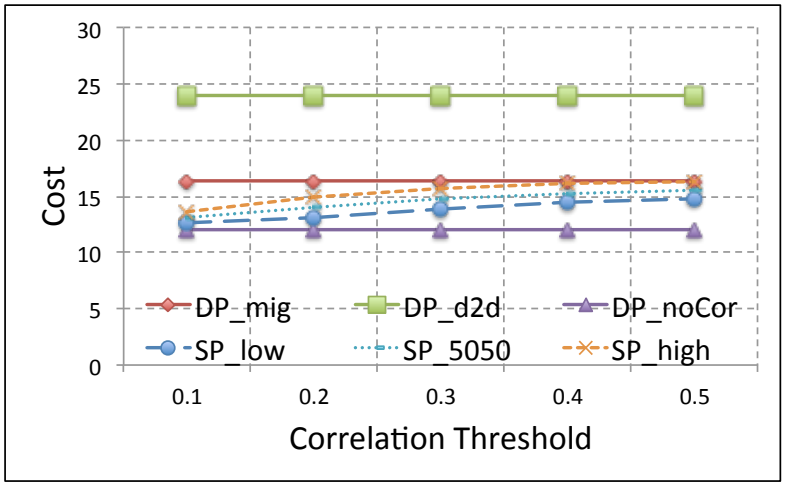

Figure 5.2 – Total Cost vs Correlation Threshold - 40 lightpath/VM requests

among these problems across all threshold values and for all three total traffic volumes considered. The threshold value is compared with the absolute difference of the service values related to their correlation. As this value increases, the cost results of the stochastic problems approach the upper bound of the deterministic problem, where migration is considered for second-stage communication. This is clear if we recall that the stochastic problems are modeled in a way that the lower cost solution is offered by enabling either VM migration or D2D path for the second-stage communication of each

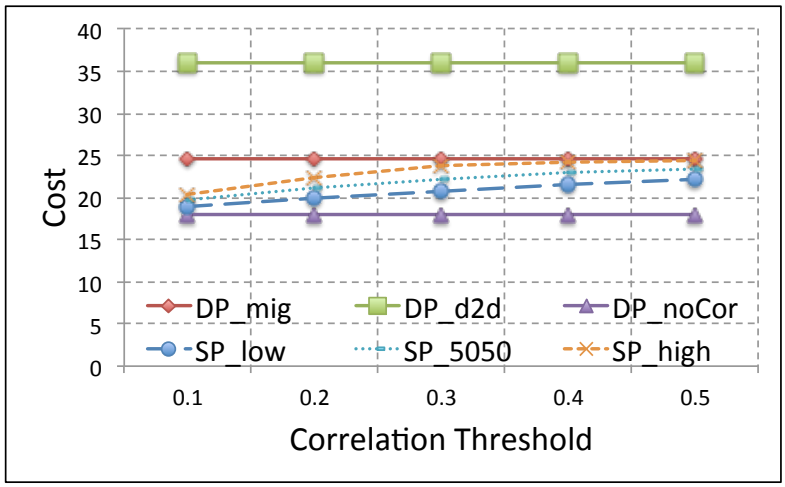

Figure 5.3 – Total Cost vs Correlation Threshold - 60 lightpath/VM requests

request. However, the relatively lower cost of the VM migration compared to the D2D path establishment results in migration for all requests, as this is the optimal choice and the reason for which the result of the DP_{mig} problem is an upper bound for the stochastic problems. In a different case where both VM migration and D2D paths would be part of the solution, the stochastic problems' cost would reside between the same lower bound, DP_{noCor} and the worst-case upper bound, DP_{d2d}. Finally, for small correlation threshold, i.e. when having a requirement for values with very small absolute difference for the service pairs obtained from the bi-variate normal distribution (a result that is already driven by the correlation coefficient), the stochastic problems' costs are close to the case where no correlation takes place, since this requirement lowers the probability of two services in a group to require communication.

For the same set of problems, threshold values and traffic volumes, Figure 5.4 - Figure 5.6 provide the results of the total network capacity in terms of the total number of wavelengths. Despite the clear differentiation of the worst-case DP_{d2d} problem results where an important amount of additional capacity is assigned due to the establishment of D2D paths, this is not the case among the rest of the problems. The DP_{noCor} is still a lower bound, but this is not clear in the results since all other problems' capacity assignment is similar and due to the small number of samples and sample size that have

an important limitation on the solution quality.

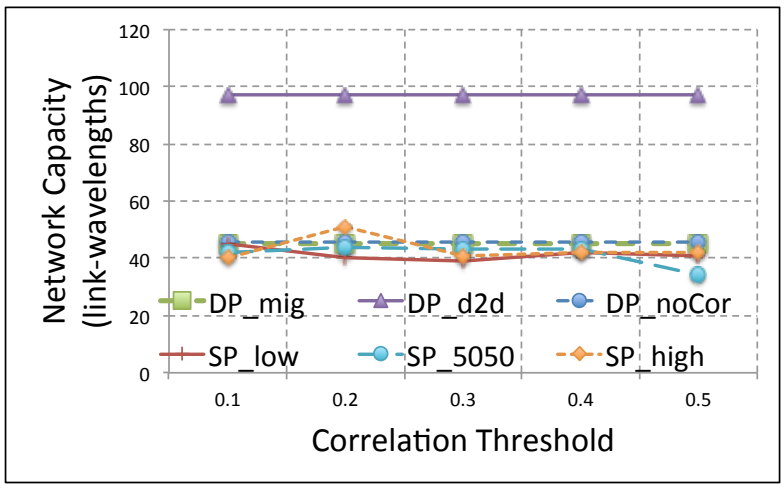

Figure 5.4 – Network Capacity vs Correlation Threshold - 20 lightpath/VM requests

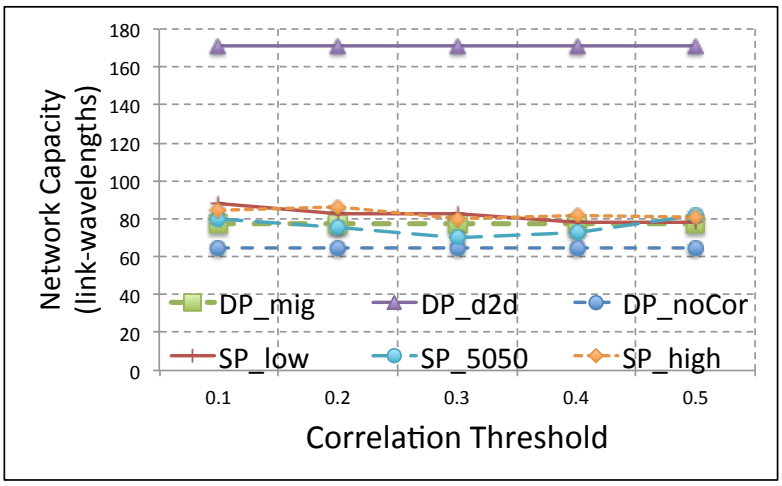

Figure 5.5 – Network Capacity vs Correlation Threshold - 40 lightpath/VM requests

The results regarding the total number of VMs deployed across the four DCs are presented in Figure 5.7 - Figure 5.9 for 20, 40 and 60 requests respectively. DP_all represents the deterministic problems DP_{mig} and DP_{d2d} since there is no differentiation on the number of VMs needed when all services are deterministically considered correlated with their pairs. This makes obvious the reason for which in all cases these results are

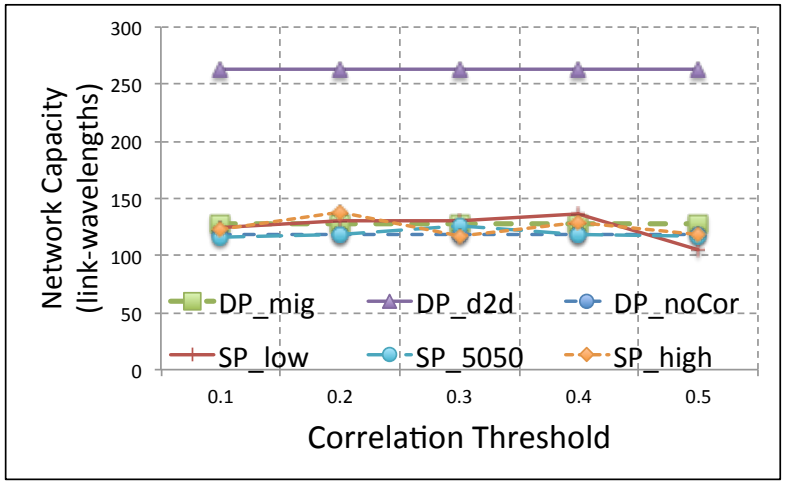

Figure 5.6 – Network Capacity vs Correlation Threshold - 60 lightpath/VM requests

fixed and equal to double the total traffic volume, since two VMs are needed for each request, one first-stage and one second-stage. The DP_{noCor} deterministic problem still provides a lower bound that is again fixed and equal to the number of requests, since no correlation is considered for any service. The stochastic problems results follow the same pattern with the cost-related results, since the number of VMs deployed gets higher when more higly correlated services are part of the total traffic requested.

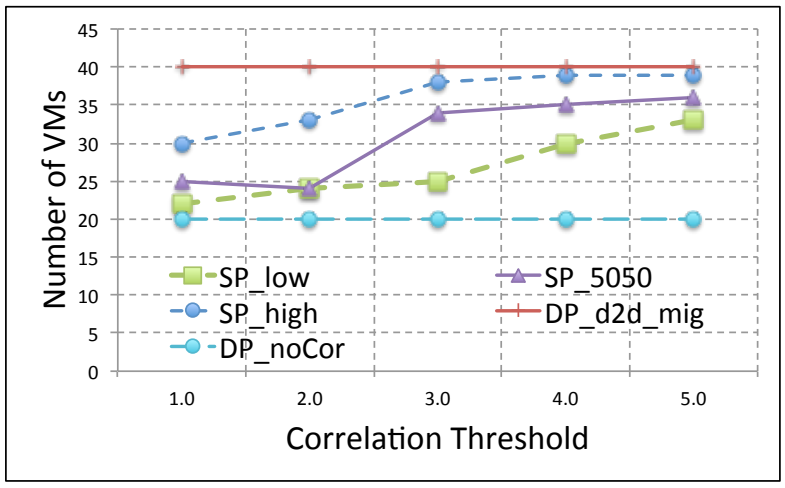

Figure 5.7 – Number of VMs vs Correlation Threshold - 20 lightpath/VM requests

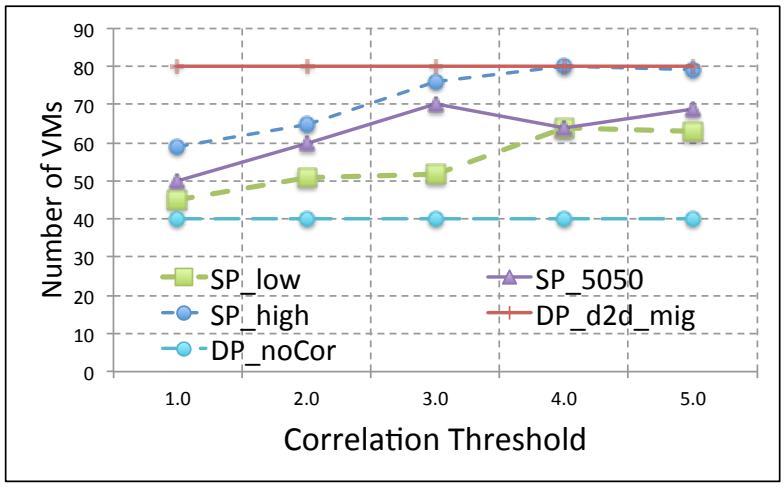

Figure 5.8 – Number of VMs vs Correlation Threshold - 40 lightpath/VM requests

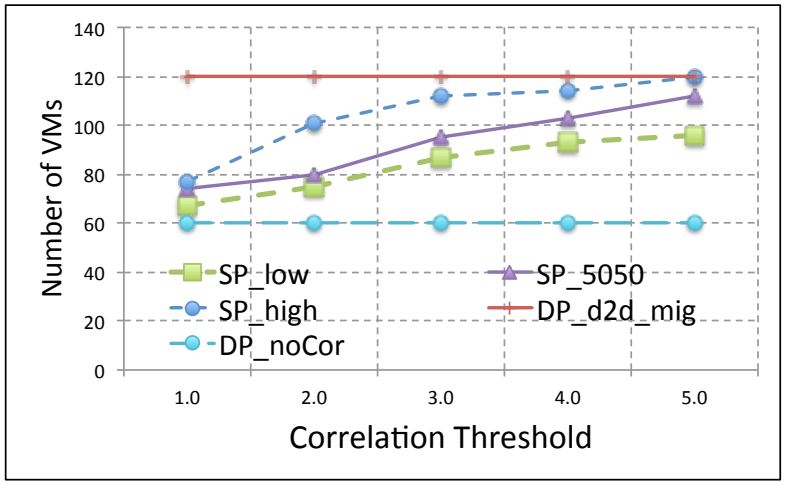

Figure 5.9 – Number of VMs vs Correlation Threshold - 60 lightpath/VM requests

To demonstrate how the solution estimate quality gets better with the number of samples, we present in Figure 5.10 the optimality gap estimate for values of M equal to 10, 20 and 50. This is based on a stochastic problem solution for 10 lightpath requests, since the increasing number of samples with a fixed sample size of 100 makes the problem very difficult to solve. We observe that if problems are solved for higher number of samples, the optimality gap estimate is subject to exponential decay.

Finally, in Figure 5.11 and Figure 5.12 we explore the same data set from a different

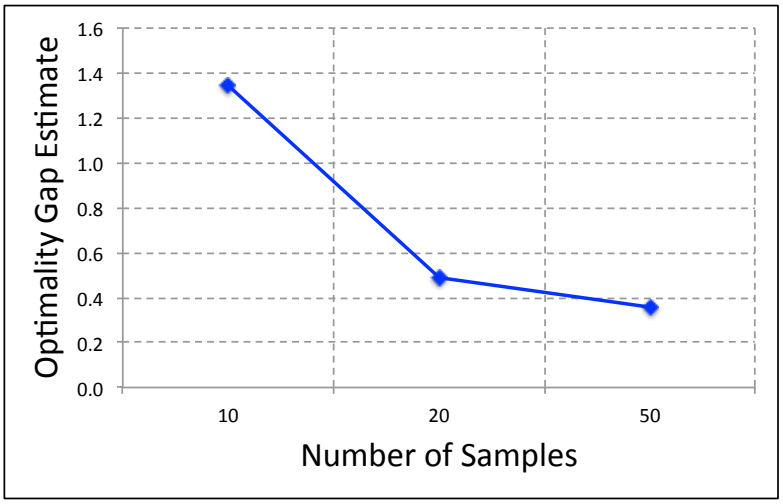

Figure 5.10 – Network Capacity vs Correlation Threshold - 20 lightpath/VM requests

point of view and present the threshold value required for different upper bounds on the total cost of the three stochstic problems. Two main observations can be made: a) Lowering the maximum cost accepted for deploying the VMs and assigning the required network capacity requires to decrease the threshold value in a different way across the three mixtures of traffic. Lowering the threshold means that we are more strict in considering when second-stage communication is required. This choice can keep the cost in the desired level but imposes an important trade-off: our estimation on the needs for required second-stage communication according to the threshold may in practice result in inadequate VM deployment and capacity placement and thus either impose a quality decrease or require re-establishment of resources within very strict time limits. b) Different threshold values are required across different traffic patterns for the same upper bound of the cost. This means that even if the estimation regarding correlation is considered to be acceptable and its trade-offs can be afforded, an adaptive way of selecting the threshold may be required according to the percentage of highly and low correlated traffic in the network. If such an approach is not available, the minimum threshold can only be applied uniformly, without considering the details of the traffic pattern. This may result in some cases in over-estimation of the required resources as

more second-stage communications will be planned and not applied when the actual services will be requested.

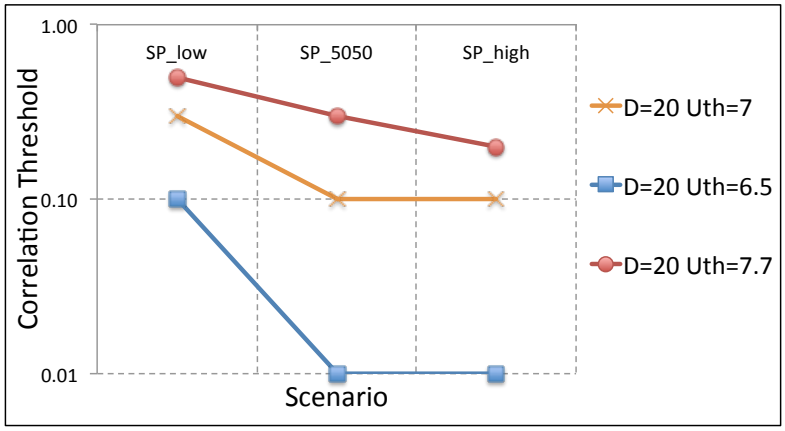

Figure 5.11 – Correlation Thresold vs Traffic Scenario - 20 lightpath/VM requests

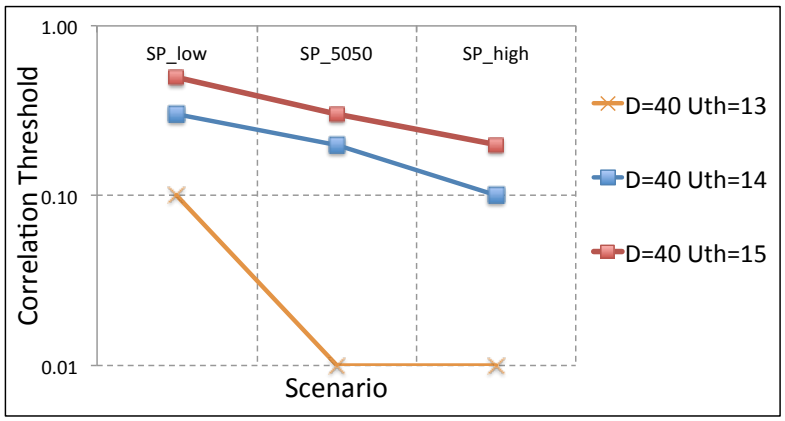

Figure 5.12 – Correlation Thresold vs Traffic Scenario - 40 lightpath/VM requests

5.5 Conclusions

In this chapter we have presented optimal VM deployment problems and solutions for minimum cost over an optical network infrastructure interconnecting a set of Data Centers. Motivated by recent studies of traffic between client nodes and Data Centers and the findings indicating correlation patterns among such cloud-based services, we formulated problems that take into account the cross-correlation between pairs of services in

the deployment process. The correlation between pairs of services has been assumed to create the need for additional inter-DC communication in order for clients to be served by the available DCs. This additional communication is modeled either through D2D paths or the VM migration process and a stochastic optimization problem has been formulated to account for the random nature of correlation and illustrate through the produced solutions the impact of handling service correlation as a random process on the total cost. Appropriate deterministic problems are formulated to produce lower and upper bounds and illustrate extreme cases that handle correlation either as non-existing or as a deterministic process that requires additional communication for every single request. The stochastic problems have been solved assuming different traffic patterns related to the percentage of highly and low correlated traffic. The results demonstrate two main findings: a) the impact that service correlation has on the additional DC and network resources required in the context of deterministic problems and b) the significant savings in cost illustrated through lower network capacity and more efficient VM deployment when correlation is handled as a random variable.

Chapter 6

Summary

This thesis focuses on service provisioning and network design problems aiming to illustrate some key problems and challenges and demonstrate efficient solutions towards core optical networks that can efficiently and effectively support the increasing volume of traffic. We have thus addressed four main areas of interest related to optical networks: a) physical layer impairments and resilence, b) resilience and energy efficiency, c) converged optical network and DC infrastructure planning and d) deployment of correlated cloud-based services over optical networks.

We have presented a performance evaluation study of an impairment aware routing algorithm in the presence of dual link failures and shared path protection. This has led to two main findings: network performance in terms of blocking and capacity utilization was related with the physical performance of the network links and the second evaluated the network performance of the different routing/sharing scheme combinations in the presence of dual-link failures. When the quality of the formed paths was taken into account in the process of routing, the physical performance of the selected paths was shown to be better, compared to the traditional routing schemes that consider only bandwidth availability. The finding was driven by the almost invariant blocking probability of all schemes for a wide loading range. The main contributor to this blocking was the dominant effect of physical impairments in the backup paths for the different simulated scenarios. The results showed a relatively constant response of the network to incoming requests for the whole loading range. Regarding the dual-link failures, although the absolute values of the average connection loss rate implied high tolerance of single-link failure-resilient networks in the presence of dual failures, the results were identified as

sufficient to violate the required 99.999% availability of carrier-grade networks, indicating the importance of spare capacity placement and/or restoration schemes against dual link failures. An overall evaluation of the results lead to the conclusion that the consideration of impairments in the routing algorithm for both primary and backup paths provides a reduced network blocking probability and thus an improved overall solution in the presence of single and dual-link failures.

Furthermore, we have evaluated the impact of all-optical wavelength conversion technology on the overall power consumption of resilient WDM optical networks. After considering the cases of no protection, 1:1 dedicated path protection and shared backup path protection schemes, we demonstrated that the shared path protection scheme achieves increased efficiency in the utilization of network resources. In combination with the energy-efficient technology of all-optical wavelength conversion, we illustrated how this scheme provides resilience to single link failures at the expense of minimal increase in the total power consumption of the network compared to the unprotected case.

Entering the field of converged optical networks and Data Center infrastructures, we presented a detailed study of planning virtual infrastructures over a physical infrastructure comprising integrated optical network and Data Center resources. The study assumed a practical VI demand model that did not support any in advance global knowledge of the VI requests. The various scenarios under study were compared with regards to power consumption, network utilization and blocking performance of the planned VIs. Under this specific VI request assumption, our results illustrated that although power consumption is an important aspect and an objective function (MJP) that optimizes the energy efficiency of the infrastructure, it may introduce inefficiencies in the utilization of network resources when the number of requests exceeds a certain level. This may in turn compromise the benefit with regards to energy efficiency, compared to what is achieved when applying an objective that minimizes the network resource utilization for this high demand levels. To overcome this inefficiency, periodic re-planning of the requests can be applied. Finally, a set of dynamic traffic provisioning results were provided through simulations illustrating that the efficient resource utilization of the second objective (MNR) introduces a penalty on the produced VIs, especially in terms of connectivity, that leads to poor request blocking performance

Lastly, we have presented optimal VM deployment problems and solutions for minimum cost over an optical network infrastructure interconnecting a set of Data Centers.

Motivated by recent studies of traffic between client nodes and Data Centers and the findings indicating correlation patterns among such cloud-based services, we formulated problems that take into account the cross-correlation between pairs of services in the deployment process. The correlation between pairs of services has been assumed to create the need for additional inter-DC communication is required in order for clients to be served by the available DCs. This additional communication is modeled either through D2D paths or the VM migration process and a stochastic optimization problem has been formulated to account for the random nature of correlation and illustrate through the produced solutions the impact of handling service correlation as a random process on the total cost. Appropriate deterministic problems are formulated to produce lower and upper bounds and illustrate extreme cases that handle correlation either as non-existing or as a deterministic process that requires additional communication for every single request. The stochastic problems have been solved assuming different traffic patterns related to the percentage of highly and low correlated traffic. The results demonstrate two main findings: a) the impact that service correlation has on the additional DC and network resources required in the context of deterministic problems and b) the significant savings in cost illustrated through lower network capacity and more efficient VM deployment when correlation is handled as a random variable.

Bibliography

[1] "User scenarios 2020 - a worldwide wireless future," L. Sørensen and K. E. Skouby, Eds. Wireless World Reseach Forum, 2009. [Online]. Available: http://www.wireless-world-research.org/fileadmin/sites/default/files/publications/Outlook/Outlook4.pdf

[2] S. Matsuoka, "Ultrahigh-speed ultrahigh-capacity transport network technology for cost-effective core and metro networks." NTT Network Innovation Laboratories, 2011. [Online]. Available: https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201108fa1.pdf&mode=show_pdf

[3] IBM, "What is big data?" [Online]. Available: http://www-01.ibm.com/software/data/bigdata/

[4] L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A break in the clouds: towards a cloud definition," *SIGCOMM Comput. Commun. Rev.*, vol. 39, no. 1, pp. 50–55, Dec. 2008. [Online]. Available: http://doi.acm.org/10.1145/1496091.1496100

[5] "Smart 2020: Enabling the low carbon economy in the information age." Global eSustainability Initiative, 2008. [Online]. Available: http://smart2020.org/_assets/files/02_Smart2020Report.pdf

[6] Y. Chen, S. Jain, V. K. Adhikari, Z.-L. Zhang, and K. Xu, "A First Look at Inter-Data Center Traffic Characteristics via Yahoo! Datasets," in *IEEE INFOCOM 2011 - IEEE Conference on Computer Communications*. IEEE, 2011, pp. 1620–1628.

[7] M. P. Jean-Philippe Vasseur and P. Demeester, *Network Recovery, Protection and Restoration of Optical, SONET-SDH, IP, and MPLS*. Morgan Kaufmann, Aug. 2004.

[8] "Introduction to dwdm technology." Cisco Systems, Inc, 2001.

[9] L. Noirie, "The road towards all-optical networks," in *Optical Fiber Communication Conference*. Optical Society of America, 2003, p. FA1. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=OFC-2003-FA1

[10] J. Berthold, A. A. M. Saleh, L. Blair, and J. M. Simmons, "Optical networking: Past, present, and future," *J. Lightwave Technol.*, vol. 26, no. 9, pp. 1104–1118, May 2008. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-26-9-1104

[11] M. J. O'Mahony, C. Politi, D. Klonidis, R. Nejabati, and D. Simeonidou, "Future optical networks," *J. Lightwave Technol.*, vol. 24, no. 12, pp. 4684–4696, Dec 2006. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-24-12-4684

[12] V. Kuroyanagi, "Photonic networking using optical add drop multiplexers and optical cross-connects," *Fujitsu Sci Tech J*, 1999.

[13] A. Tzanakaki, I. Zacharopoulos, and I. Tomkos, "Broadband building blocks [optical networks]," *Circuits and Devices Magazine, IEEE*, vol. 20, no. 2, pp. 32 – 37, mar-apr 2004.

[14] G. N. Rouskas and H. G. Perros, "A tutorial on optical networks," in *Advanced lectures in Networking*, 2002, pp. 155–193.

[15] "Modernizing the core with otn switching & 40/100g transport." heavyreading.com, 2010.

[16] A. Dutta, N. Dutta, and M. Fujiwara, *WDM Technologies: Optical Networks*. Academic Press, 2004.

[17] K. Katrinis and A. Tzanakaki, "On the dimensioning of wdm optical networks with impairment-aware regeneration," *Networking, IEEE/ACM Transactions on*, vol. 19, no. 3, pp. 735 –746, june 2011.

[18] G. Markidis, S. Sygletos, A. Tzanakaki, and I. Tomkos, "Impairment aware based routing and wavelength assignment in transparent long haul networks," in *Optical Network Design and Modeling*, ser. Lecture Notes in Computer Science, I. Tomkos, F. Neri, J. Solé Pareta, X. Masip Bruin, and S. Sánchez Lopez,

Eds. Springer Berlin Heidelberg, 2007, vol. 4534, pp. 48–57. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-72731-6_6

[19] G. Markidis and A. Tzanakaki, "Routing and Wavelength Assignment Algorithms in Survivable WDM Networks Under Physical Layer Constraints," in *5th International Conference on Broadband Communications, Networks and Systems, 2008 (BROADNETS 2008)*, 2008.

[20] G. Shen and R. S. Tucker, "Energy-Minimized Design for IP Over WDM Networks," *Journal of Optical Communications and Networking*, vol. 1, no. 1, pp. 176–186, 2009.

[21] A. Muhammad, P. Monti, I. Cerutti, L. Wosinska, P. Castoldi, and A. Tzanakaki, "Energy-Efficient WDM Network Planning with Dedicated Protection Resources in Sleep Mode," *2010 IEEE Global Telecommunications Conference (GLOBECOM 2010)*, pp. 1–5, 2010.

[22] D. Bertsekas, *Network optimization: continuous and discrete methods*, ser. Optimization and neural computation series. Athena Scientific, 1998. [Online]. Available: http://books.google.gr/books?id=afYYAQAAIAAJ

[23] E. Escalona, S. Peng, R. Nejabati, D. Simeonidou, J. Garcia-Espin, J. Ferrer, S. Figuerola, G. Landi, N. Ciulli, J. Jimenez, B. Belter, Y. Demechenko, C. de Laat, X. Chen, A. Yukan, S. Soudan, P. Vicat-Blanc, J. Buysse, M. De Leenheer, C. Develder, A. Tzanakaki, P. Robinson, M. Brogle, and T. Bohnert, "Geysers: A novel architecture for virtualization and co-provisioning of dynamic optical networks and it services," in *Future Network Mobile Summit (FutureNetw), 2011*, june 2011, pp. 1 –8.

[24] W. D. Grover, *Mesh-based Survivable Transport Networks: Options and Strategies for Optical, MPLS, SONET and ATM Networking.* Upper Saddle River, NJ, USA: Prentice Hall PTR, 2003.

[25] B. Verweij, S. Ahmed, A. Kleywegt, G. Nemhauser, and A. Shapiro, "The sample average approximation method applied to stochastic routing problems: A computational study," *Computational Optimization and Applications*, vol. 24, pp. 289–333, 2003.

[26] J.-F. Labourdette, E. Bouillet, R. Ramamurthy, and G. Ellinas, *Path Routing in Mesh Optical Networks.* John Wiley & Sons, 2006.

[27] W. D. Grover, M. Clouqueur, and T. Bach, "Quantifying and managing the influence of maintenance actions on the survivability of mesh-restorable networks," in *Proceedings 17th Annual National Fiber Optic Engineers Conference (NFOEC 2001*, 2001, pp. 8–12.

[28] W. He and A. Somani, "Path-based protection for surviving double-link failures in mesh-restorable optical networks," in *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, vol. 5, dec. 2003, pp. 2558 – 2563 vol.5.

[29] L. Guo, L. Li, J. Cao, H. Yu, and X. Wei, "On finding feasible solutions with shared backup resources for surviving double-link failures in path-protected wdm mesh networks," *Journal of Lightwave Technology*, vol. 25, no. 1, pp. 287 –296, jan. 2007.

[30] J. Doucette and M. Clouqueur, *On the availability and capacity requirements of shared backup path-protected mesh networks.* Optical Networks Magazine, 2003.

[31] S. Ramasubramanian and A. Chandak, "Dual-link failure resiliency through backup link mutual exclusion," *Networking, IEEE/ACM Transactions on*, vol. 16, no. 1, pp. 157 –169, feb. 2008.

[32] S. Ahuja and S. Ramasubramanian, "Enhancing Robustness Under Dual-Link Failures," in *Proceedings of 16th International Conference on Computer Communications and Networks, 2007*, 2007.

[33] Y. Kavian, R. Rejeb, and O. Strobel, "Dual-link failure covering in dwdm optical networks using genetic algorithms," in *Transparent Optical Networks (ICTON), 2010 12th International Conference on*, 27 2010-july 1 2010, pp. 1 –4.

[34] J. Zhang and B. Mukheriee, "A review of fault management in wdm mesh networks: basic concepts and research challenges," *Network, IEEE*, vol. 18, no. 2, pp. 41 – 48, mar-apr 2004.

[35] C. Politi, V. Anagnostopoulos, C. Matrakidis, and A. Stavdas, "Physical layer impairment aware routing algorithms based on analytically calculated q-factor," in

Optical Fiber Communication Conference, 2006 and the 2006 National Fiber Optic Engineers Conference. OFC 2006, march 2006, p. 3 pp.

[36] C. Saradhi and S. Subramaniam, "Physical layer impairment aware routing (pliar) in wdm optical networks: issues and challenges," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 4, pp. 109 –130, quarter 2009.

[37] A. G. Rahbar, "Review of dynamic impairment-aware routing and wavelength assignment techniques in all-optical wavelength-routed networks," *Communications Surveys Tutorials, IEEE*, vol. 14, no. 4, pp. 1065 –1089, quarter 2012.

[38] A. Jirattigalachote, Y. Yamada, C. Cavdar, P. Monti, L. Wosinska, H. Hasegawa, and K. ichi Sato, "Impairment-aware routing and waveband assignment for efficient optical transport networks," in *Optical Fiber Communication Conference*. Optical Society of America, 2012, p. OW3A.2. [Online]. Available: http://www.opticsinfobase.org/abstract.cfm?URI=OFC-2012-OW3A.2

[39] P. Batchelor, B. Daino, P. Heinzmann, D. Hjelme, R. Inkret, H. Jager, M. Joindot, A. Kuchar, E. Coquil, P. Leuthold, G. Marchis, F. Matera, B. Mikac, H.-P. Nolting, J. Spath, F. Tillerot, B. Caenegem, N. Wauters, and C. Weinert, "Study on the implementation of optical transparent transport networks in the european environment—results of the research project cost 239," *Photonic Network Communications*, vol. 2, pp. 15–32, 2000. [Online]. Available: http://dx.doi.org/10.1023/A%3A1010050906938

[40] K. C. Claffy, H.-W. Braun, and G. C. Polyzos, "Tracking long-term growth of the nsfnet," *Commun. ACM*, vol. 37, no. 8, pp. 34–45, Aug. 1994. [Online]. Available: http://doi.acm.org/10.1145/179606.179616

[41] M. Clouqueur and W. D. Grover, "Availability Analysis of Span-Restorable Mesh Networks," *IEEE Journal on Selected Areas in Communications*, pp. 1–12, Dec. 2002.

[42] D. A. Schupke, A. Autenrieth, and T. Fischer, "Survivability of multiple fiber duct failures," in *Third International Workshop on the Design of Reliable Communication Networks (DRCN), year = 2001, pages = 7–10.*

[43] S. Kim and S. Lumetta, "Evaluation of protection reconfiguration for multiple failures in wdm mesh networks," in *Optical Fiber Communications Conference, 2003. OFC 2003*, march 2003, pp. 210 – 211 vol.1.

[44] C. Assi, W. Huo, and A. Shami, "Impact of resource sharability on dual failure restorability in optical mesh networks," in *Proceedings of the 4th IFIP-TC6 international conference on Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communication Systems*, ser. NETWORKING'05. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 792–803. [Online]. Available: http://dx.doi.org/10.1007/11422778_64

[45] D. Schupke and R. Prinz, "Capacity efficiency and restorability of path protection and rerouting in WDM networks subject to dual failures," *Photonic Network Communications*, vol. 8, no. 2, pp. 191–207, 2004.

[46] B. Mukherjee and Y. Huang, "Impairment-aware routing in wavelength-routed optical networks," *The 17th Annual Meeting of the IEEE Lasers and Electro-Optics Society, 2004. LEOS 2004.*, 2004.

[47] I. Tomkos, S. Sygletos, A. Tzanakaki, and G. Markidis, "Impairment Constraint Based Routing in Mesh Optical Networks," in *Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference (2007), paper OWR1.* Optical Society of America, Mar. 2007.

[48] S. Pachnicke and N. Luck, "Online Physical-Layer Impairment-Aware Routing with Quality of Transmission Constraints in Translucent Optical Networks," in *11th International Conference on Optical Networks*, 2009, pp. 1–4.

[49] M. Pickavet, W. Vereecken, S. Demeyer, P. Audenaert, B. Vermeulen, C. Develder, D. Colle, B. Dhoedt, and P. Demeester, "Worldwide Energy Needs for leT: the Rise ofPower-Aware Networking," in *2008 2nd International Symposium on Advanced Networks and Telecommunication Systems (ANTS).* IEEE, pp. 1–3.

[50] J. Baliga, R. Ayre, K. Hinton, W. V. Sorin, and R. S. Tucker, "Energy consumption in optical ip networks," *J. Lightwave Technol.*, vol. 27, no. 13, pp. 2391–2403, Jul 2009. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-27-13-2391

[51] J. Chabarek, J. Sommers, P. Barford, C. Estan, D. Tsiang, and S. Wright, "Power Awareness in Network Design and Routing," in *IEEE INFOCOM 2008 - IEEE Conference on Computer Communications*. IEEE, 2008, pp. 457–465.

[52] L. Chiaraviglio, M. Mellia, and F. Neri, "Reducing Power Consumption in Backbone Networks," in *ICC 2009 - 2009 IEEE International Conference on Communications*. IEEE, 2009, pp. 1–6.

[53] M. Murakami, "Analyzing Power Consumption in Optical Cross-connect Equipment for Future Large-Capacity Optical Networks," *Journal of Networks*, vol. 5, no. 11, Nov. 2010.

[54] http://www.google.com/about/datacenters/locations/index.html.

[55] T. Benson, A. Akella, and D. A. Maltz, "Network traffic characteristics of data centers in the wild," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, ser. IMC '10. New York, NY, USA: ACM, 2010, pp. 267–280. [Online]. Available: http://doi.acm.org/10.1145/1879141.1879175

[56] Y. Zhang, P. Chowdhury, M. Tornatore, and B. Mukherjee, "Energy Efficiency in Telecom Optical Networks," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 4, pp. 441–458, 2010.

[57] C. Develder, M. De Leenheer, B. Dhoedt, M. Pickavet, D. Colle, F. De Turck, and P. Demeester, "Optical networks for grid and cloud computing applications," *Proceedings of the IEEE*, vol. 100, no. 5, pp. 1149 –1167, may 2012.

[58] M. Ritter, "Virtualized Optical Networks for Sustainable Cloud Services," ADVA Optical Networking, Tech. Rep., Jan. 2010.

[59] X. Liu, C. Qiao, W. Wei, X. Yu, T. Wang, W. Hu, W. Guo, and M.-Y. Wu, "Task scheduling and lightpath establishment in optical grids," *J. Lightwave Technol.*, vol. 27, no. 12, pp. 1796–1805, Jun 2009. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-27-12-1796

[60] J. KJ (Ken) Salchow, "Data Centre Consolidation: Know Where You're Going and Why," F5 Networks, Inc, Tech. Rep., Oct. 2010.

[61] A. Tzanakaki, M. Anastasopoulos, K. Georgakilas, J. Buysse, M. De Leenheer, C. Develder, S. Peng, R. Nejabati, E. Escalona, D. Simeonidou, N. Ciulli, G. Landi, M. Brogle, A. Manfredi, E. Lopez, J. F. Riera, J. A. Garcia-Espin, P. Donadio, G. Parladori, and J. Jimenez, "2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)," in *IEEE INFOCOM 2011 - IEEE Conference on Computer Communications Workshops*. IEEE, 2011, pp. 343–348.

[62] M. Pickavet, W. Vereecken, S. Demeyer, P. Audenaert, B. Vermeulen, C. Develder, D. Colle, B. Dhoedt, and P. Demeester, "Worldwide energy needs for ict: The rise of power-aware networking," in *Advanced Networks and Telecommunication Systems, 2008. ANTS '08. 2nd International Symposium on*, dec. 2008, pp. 1 –3.

[63] L. Chiaraviglio, M. Mellia, and F. Neri, "Reducing power consumption in backbone networks," in *Communications, 2009. ICC '09. IEEE International Conference on*, june 2009, pp. 1 –6.

[64] J. Baliga, R. Ayre, K. Hinton, and R. Tucker, "Green cloud computing: Balancing energy in processing, storage, and transport," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 149 –167, jan. 2011.

[65] J. Buysse, K. Georgakilas, A. Tzanakaki, M. De Leenheer, B. Dhoedt, C. Develder, and P. Demeester, "Calculating the Minimum Bounds of Energy Consumption for Cloud Networks," in *ICCCN*, Feb. 2011, pp. 1–8.

[66] A. Tzanakaki, M. P. Anastasopoulos, K. Georgakilas, and D. Simeonidou, "Energy Aware Planning of Multiple Virtual Infrastructures Over Converged Optical Network and IT Physical Resources," in *ECOC*, Apr. 2011, pp. 1–3.

[67] M. Jinno and Y. Tsukishima, "Virtualized optical network (von) for agile cloud computing environment," in *Optical Fiber Communication - incudes post deadline papers, 2009. OFC 2009. Conference on*, march 2009, pp. 1 –3.

[68] B. Stiller, T. Bocek, F. Hecht, G. Machado, P. Racz, and M. Waldburger, "Oracle Exalogic Elastic Cloud X2-2 Data Sheet," Tech. Rep., Mar. 2011.

[69] Oracle. http://www.oracle.com/technetwork/database/exadata/dbmachine-x2-8-datasheet-173705.pdf .

[70] S. Aleksić, "Analysis of Power Consumption in Future High-Capacity Network Nodes," *Journal of Optical Communications and Networking*, vol. 1, no. 3, p. 245, 2009.

[71] A. Tzanakaki, K. Katrinis, T. Politi, A. Stavdas, M. Pickavet, P. Van Daele, D. Simeonidou, M. J. O'Mahony, S. Aleksić, L. Wosinska, and P. Monti, "Dimensioning the Future Pan-European Optical Network With Energy Efficiency Considerations," *Journal of Optical Communications and Networking*, vol. 3, no. 4, p. 272, 2011.

[72] J. Y. Yen, "Finding the k shortest loopless paths in a network," *Management Science*, vol. 17, no. 11, pp. 712–716, 1971. [Online]. Available: http://mansci.journal.informs.org/content/17/11/712.abstract

[73] K. N. Georgakilas, K. Katrinis, A. Tzanakaki, and O. B. Madsen, "Performance Evaluation of Impairment-Aware Routing Under Single- and Double-Link Failures," *Journal of Optical Communications and Networking*, vol. 2, no. 8, p. 633, 2010.

[74] R. Buyya, "Market-oriented cloud computing: Vision, hype, and reality of delivering computing as the 5th utility," in *Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on*, may 2009, p. 1.

[75] E. K. Galen Gruman. What cloud computing really means — Cloud Computing - InfoWorld. [Online]. Available: http://www.infoworld.com/d/cloud-computing/what-cloud-computing-really-means-031

[76] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of virtualization," *SIGOPS Oper. Syst. Rev.*, vol. 37, no. 5, pp. 164–177, Oct. 2003. [Online]. Available: http://doi.acm.org/10.1145/1165389.945462

[77] W. Voorsluys, J. Broberg, S. Venugopal, and R. Buyya, "Cost of Virtual Machine Live Migration in Clouds: A Performance Evaluation," in *Lecture Notes in Computer Science*, M. Jaatun, G. Zhao, and C. Rong, Eds. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2009, pp. 254–265.

[78] M. Zhao and R. J. Figueiredo, "Experimental study of virtual machine migration in support of reservation of cluster resources," in *Proceedings of the 2nd*

international workshop on Virtualization technology in distributed computing, ser. VTDC '07. New York, NY, USA: ACM, 2007, pp. 5:1–5:8. [Online]. Available: http://doi.acm.org/10.1145/1408654.1408659

[79] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield, "Live migration of virtual machines," in *Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation - Volume 2*, ser. NSDI'05. Berkeley, CA, USA: USENIX Association, 2005, pp. 273–286. [Online]. Available: http://dl.acm.org/citation.cfm?id=1251203.1251223

[80] F. Travostino, P. Daspit, L. Gommans, C. Jog, C. de Laat, J. Mambretti, I. Monga, B. van Oudenaarde, S. Raghunath, and P. Y. Wang, "Seamless live migration of virtual machines over the man/wan," *Future Gener. Comput. Syst.*, vol. 22, no. 8, pp. 901–907, Oct. 2006. [Online]. Available: http://dx.doi.org/10.1016/j.future.2006.03.007

[81] Q. Zhang, J. Xiao, E. Gürses, M. Karsten, and R. Boutaba, "Dynamic Service Placement in Shared Service Hosting Infrastructures," *NETWORKING 2010*, 2010.

[82] Y. Kang, Z. Zheng, and M. R. Lyu, "A Latency-Aware Co-Deployment Mechanism for Cloud-Based Services," in *2012 IEEE 5th International Conference on Cloud Computing (CLOUD)*. IEEE, pp. 630–637.

[83] M. Steiner, B. G. Gaglianello, V. Gurbani, V. Hilt, W. D. Roome, M. Scharf, and T. Voith, "Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication - SIGCOMM '12," in *the ACM SIGCOMM 2012 conference*. New York, New York, USA: ACM Press, 2012, p. 73.

[84] L. F. Birge, John R., *Introduction to Stochastic Programming*. Springer Series in Operations Research and Financial Engineering, 2011.

[85] G. Andreatta and L. Romeo, "Stochastic shortest paths with recourse," *Networks*, vol. 18, no. 3, pp. 193–204, 1988.

[86] A. Marchetti Spaccamela, A. H. G. Rinnooy Kan, and L. Stougie, "Hierarchical vehicle routing problems," *Networks*, vol. 14, no. 4, pp. 571–586, 1984.

[87] G. Laporte and F. V. Louveaux, "The integer L-shaped method for stochastic integer programs with complete recourse," *Operations Research Letters*, vol. 13, no. 3, pp. 133–142, Apr. 1993.